

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 November 1999 (11.11.1999)

PCT

(10) International Publication Number
WO 99/057280 A3

(51) International Patent Classification⁶: **C12N 15/31**,
C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N
33/50

(21) International Application Number: PCT/US99/09346

(22) International Filing Date: 30 April 1999 (30.04.1999)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/083,758	1 May 1998 (01.05.1998)	US
60/094,869	31 July 1998 (31.07.1998)	US
60/098,994	2 September 1998 (02.09.1998)	US
60/099,062	2 September 1998 (02.09.1998)	US
60/103,749	9 October 1998 (09.10.1998)	US
60/103,794	9 October 1998 (09.10.1998)	US
60/103,796	9 October 1998 (09.10.1998)	US
60/121,528	25 February 1999 (25.02.1999)	US

(71) Applicants (for all designated States except US): **CHIRON CORPORATION** [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). **THE INSTITUTE FOR GENOMIC RESEARCH** [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **FRASER, Claire** [US/US]; The Institute for Genomic Research, 9212 Medical Center Drive, Rockville, MD 20850 (US). **GALLEOTTI, Cesira** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **GRANDI, Guido** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **HICKEY, Erin** [US/US]; The Institute for Genomic Research, 9212 Medical Center Drive, Rockville, MD 20850 (US). **MASIGNANI, Vega** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **MORA, Marirosa** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **PETERSEN, Jeremy** [US/US]; The Institute for Genomic Research, 9212 Medical Center Drive, Rockville, MD 20850 (US). **PIZZA, Mariagrazia** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **RAPPUOLI, Rino** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100

Siena (IT). **RATTI, Giulio** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **SCALATO, Enzo** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **SCARSELLI, Maria** [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). **TETTELIN, Herve** [US/US]; The Institute for Genomic Research, 9212 Medical Center Drive, Rockville, MD 20850 (US). **VENTER, J., Craig** [US/US]; The Institute for Genomic Research, 9212 Medical Center Drive, Rockville, MD 20850 (US).

(74) Agent: **HARBIN, Alisa, A.**; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
24 August 2000

(48) Date of publication of this corrected version:
29 August 2002

(15) Information about Correction:
see PCT Gazette No. 35/2002 of 29 August 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

(57) Abstract: The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.



WO 99/057280 A3

NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25, 1999).

This invention relates to antigens from the bacterial species: *Neisseria meningitidis* and *Neisseria gonorrhoeae*.

BACKGROUND

Neisseria meningitidis is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to *N. gonorrhoea*, although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

N. meningitidis causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman *et al.* (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. *JAMA* 275(19):1499-1503; Schuchat *et al* (1997) Bacterial Meningitis in the United States in 1995. *N Engl J Med* 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against *Haemophilus influenzae*, *N. meningitidis* is the major cause of bacterial meningitis at all ages in the United States (Schuchat *et al* (1997) *supra*).

Based on the organism's capsular polysaccharide, 12 serogroups of *N. meningitidis* have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [*eg.* Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: *New Generation Vaccines, supra*, pp. 469-488; Lieberman *et al* (1996) *supra*; Costantino *et al* (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. *Vaccine* 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of $\alpha(2-8)$ -linked *N*-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the *N*-acetyl groups with *N*-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? *Clin Microbiol Rev* 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (*eg.* Poolman JT (1992) Development of a meningococcal vaccine. *Infect. Agents Dis.* 4:13-28).

Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (eg. Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonococcal genes and proteins (e.g. EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic *Neisseriae* including *Neisseria meningitidis* or *Neisseria gonorrhoeae*. Those sequences specific to *N. meningitidis* or *N. gonorrhoeae* that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters: gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples. The fragments should comprise at least *n* consecutive amino acids from the sequences and, depending on the particular sequence, *n* is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (eg. recombinant expression, purification from cell culture, chemical synthesis *etc.*) and in various forms (eg. native, fusions *etc.*). They are preferably prepared in substantially pure or isolated form (*ie.* substantially free from other *N. meningitidis* or *N. gonorrhoeae* host cell proteins)

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the *N. meningitidis* nucleotide sequences and *N. gonorrhoeae* nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least *n* consecutive nucleotides from the *N. meningitidis* sequences or *N. gonorrhoeae* sequences and depending on the particular sequence, *n* is 10 or more (eg 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself *etc.*) and can take various forms (*eg.* single stranded, double stranded, vectors, probes *etc.*).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) *etc.*

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (*eg.* expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (*eg.* as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as *N. gonorrhoeae*) but are preferably *N. meningitidis*, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Methodology - Summary of standard procedures and techniques.

General

This invention provides *Neisseria meningitidis* menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook *Molecular Cloning: A Laboratory Manual, Second Edition* (1989); *DNA Cloning, Volumes I and ii* (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & S.J. Higgins eds. 1984); *Animal Cell Culture* (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the *Methods in Enzymology* series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), and *Handbook of Experimental Immunology, Volumes I-IV* (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.

All publications, patents, and patent applications cited herein are incorporated in full by reference.

Expression systems

The *Neisseria menB* nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In *Molecular Cloning: A Laboratory Manual*, 2nd ed.).

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promoters may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) *Science* 236:1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) *EMBO J.* 4:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) *Proc. Natl. Acad. Sci.* 79:6777) and from human cytomegalovirus (Boshart et al. (1985) *Cell* 41:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) *Trends Genet.* 2:215; Maniatis et al. (1987) *Science* 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) *Cell* 41:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M.

Glover); Proudfoot (1989) *Trends Biochem. Sci.* 14:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) *Cell* 23:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) *Mol. Cell. Biol.* 9:946) and pHEBO (Shimizu et al. (1986) *Mol. Cell. Biol.* 6:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, *Gibberellins*: in: *Advanced Plant Physiology*, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Rept.*, 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl. Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its own set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) 17:31).

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) *Ann. Rev. Microbiol.*, 42:177) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO

Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), *J. Gen. Virol.* 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha) α -interferon, Maeda et al., (1985), *Nature* 315:592; human gastrin-releasing peptide, Lebacqz-Verheyden et al., (1988), *Molec. Cell. Biol.* 8:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; mouse IL-3, (Miyajima et al., (1987) *Gene* 58:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith *supra*; Ju et al. (1987); Smith et al., *Mol. Cell. Biol.* (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), *Bioessays* 4:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 μm in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. *Current Protocols in Microbiology* Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, *supra*; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, *inter alia*: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) *J. Virol.* 56:153; Wright (1986) *Nature* 321:718; Smith et al., (1983) *Mol. Cell. Biol.* 3:2156; and see generally, Fraser, et al. (1989) *In Vitro Cell. Dev. Biol.* 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith *supra*.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in *Escherichia coli* (*E. coli*) (Raibaud *et al.* (1984) *Annu. Rev. Genet.* 18:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) (Chang *et al.* (1977) *Nature* 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) (Goeddel *et al.* (1980) *Nuc. Acids Res.* 8:4057; Yelverton *et al.* (1981) *Nucl. Acids Res.* 9:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The beta-lactamase (*bla*) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)), bacteriophage lambda PL (Shimatake *et al.* (1981) *Nature* 292:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene* 25:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci.* 80:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol.* 189:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci.* 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine *et al.* (1975) *Nature* 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of *E. coli* 16S rRNA (Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberger)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook *et al.* (1989) "Expression of cloned genes in *Escherichia coli*." In *Molecular Cloning: A Laboratory Manual*).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai *et al.* (1984) *Nature* 309:810). Fusion proteins can also be made with sequences from the *lacZ* (Jia *et al.* (1987) *Gene* 60:197), *trpE* (Allen *et al.* (1987) *J. Biotechnol.* 5:93; Makoff *et al.* (1989) *J. Gen. Microbiol.* 135:11), and *Chey* (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller *et al.* (1989) *Bio/Technology* 7:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghrayeb *et al.* (1984) *EMBO J.* 3:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci.* 82:7212). As an additional example, the signal sequence of the alpha-amylase gene from various *Bacillus* strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various *Bacillus* strains integrate into the *Bacillus* chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies *et al.* (1978) *Annu. Rev. Microbiol.* 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: *Bacillus subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), *Escherichia coli* (Shimatake *et al.* (1981) *Nature* 292:128; Amann *et al.* (1985) *Gene* 40:183; Studier *et al.* (1986) *J. Mol. Biol.* 189:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), *Streptococcus cremoris* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655); *Streptococcus lividans* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655), *Streptomyces lividans* (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with CaCl_2 or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of *Bacillus*: Masson *et al.* (1989) *FEMS Microbiol. Lett.* 60:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of *Campylobacter*: Miller *et al.* (1988) *Proc. Natl. Acad. Sci.* 85:856; and Wang *et al.* (1990) *J. Bacteriol.* 172:949; use of *Escherichia coli*: Cohen *et al.* (1973) *Proc. Natl. Acad. Sci.* 69:2110; Dower *et al.* (1988) *Nucleic Acids Res.* 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol.* 53:159; Taketo (1988) *Biochim. Biophys. Acta* 949:318; use of *Lactobacillus*: Chassy *et al.* (1987) *FEMS Microbiol. Lett.* 44:173; use of *Pseudomonas*: Fiedler *et al.* (1988) *Anal. Biochem* 170:38; use of *Staphylococcus*: Augustin *et al.* (1990) *FEMS Microbiol. Lett.* 66:203; use of *Streptococcus*: Barany *et al.* (1980) *J. Bacteriol.* 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun.* 32:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology* 1:412.

v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.

Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA* 77:1078; Henikoff *et al.* (1981) *Nature* 283:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol.* 96:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene* 11:163; Panthier *et al.* (1980) *Curr. Genet.* 2:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator

sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene* 8:17-24), pCI/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci. USA* 81:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol.* 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake *et al.*, *supra*.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol.* 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al.*, *supra*. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol. Rev.* 51:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, *inter alia*, the following yeasts: *Candida albicans* (Kurtz, *et al.* (1986) *Mol. Cell. Biol.* 6:142); *Candida maltosa* (Kunze, *et al.* (1985) *J. Basic Microbiol.* 25:141); *Hansenula polymorpha* (Gleeson, *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302); *Kluyveromyces fragilis* (Das, *et al.* (1984) *J. Bacteriol.* 158:1165); *Kluyveromyces lactis* (De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:737; Van den Berg *et al.* (1990) *Bio/Technology* 8:135); *Pichia guilliermondii* (Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141); *Pichia pastoris* (Cregg, *et al.* (1985) *Mol. Cell. Biol.* 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163); *Schizosaccharomyces pombe* (Beach and Nurse (1981) *Nature* 300:706); and *Yarrowia lipolytica* (Davidow, *et al.* (1985) *Curr. Genet.* 10:380471 Gaillardin, *et al.* (1985) *Curr. Genet.* 10:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz *et al.* (1986) *Mol. Cell. Biol.* 6:142; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; *Candida*]; [Gleeson *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302; *Hansenula*]; [Das *et al.* (1984) *J.*

Bacteriol. 158:1165; De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:1165; Van den Berg *et al.* (1990) *Bio/Technology* 8:135; Kluyveromyces]; [Cregg *et al.* (1985) *Mol. Cell. Biol.* 5:3376; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; *Pichia*]; [Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163 *Saccharomyces*]; [Beach and Nurse (1981) *Nature* 300:706; *Schizosaccharomyces*]; [Davidow *et al.* (1985) *Curr. Genet.* 10:39; Gaillardin *et al.* (1985) *Curr. Genet.* 10:49; *Yarrowia*].

Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" *Neisseria* amino acid fragment or protein is one that is present in a particular *Neisseria* protein in at least x% of *Neisseria*. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all *Neisseria*). In order to determine whether an amino acid is "conserved" in a particular *Neisseria* protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different *Neisseria* (a reference population). The reference population may include a number of different *Neisseria* species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common *Neisseria* strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a *Neisseria* sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An “origin of replication” is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A “mutant” sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an “allelic variant” of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

Antibodies

As used herein, the term “antibody” refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An “antibody combining site” is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. “Antibody” includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying *Neisseria menB* proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments. Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (*Nature* (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly ^{32}P and ^{125}I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example, ^{125}I may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with ^{125}I , or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance

or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or

liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) RibiTM adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as StimulonTM (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648).

Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence in vivo can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; and Kaplitt (1994) *Nature Genetics* 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) *J. Virol.* 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) *J. Virol.* 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci* 81:6349; and Miller (1990) *Human Gene Therapy* 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) *Hum. Gene Ther.* 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) *Gene* 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) *J. Virol.* 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) *Human Gene Therapy* 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) *Science* 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995, WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukaryotic layered expression systems are also useful for expressing the nucleic acids of the invention. See WO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, *Nature* 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and *Nature* (1979) 277:108); human

immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Trinita virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) *J. Biol. Chem.* 262:4429-4432, insulin as described in Hucked (1990) *Biochem Pharmacol* 40:253-263, galactose as described in Plank (1992) *Bioconjugate Chem* 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; in WO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, *Biochemistry*, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) *Biochem Biophys Acta* 600:1; Bayer (1979) *Biochem Biophys Acta* 550:464; Rivnay (1987) *Meth Enzymol* 149:119; Wang (1987) *Proc Natl Acad Sci* 84:7851; Plant (1989) *Anal Biochem* 176:420.

A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

Delivery Methods

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered *ex vivo*, to cells derived from the subject; or (3) *in vitro* for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in *ex vivo* applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Polynucleotide and polypeptide pharmaceutical compositions

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

A. Polypeptides

One example are polypeptides which include, without limitation: asialoglycoprotein (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII.

B.Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

C.Polyalkylenes, Polysaccharides, etc.

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethylene glycol. In addition, mono-, di-, or polysaccharides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

D.Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner *supra*). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boehringer). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, eg. Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See eg. Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; Papahadjopoulos (1975) *Biochim. Biophys. Acta* 394:483; Wilson (1979) *Cell* 17:77; Deamer & Bangham (1976) *Biochim. Biophys. Acta* 443:629; Ostro (1977) *Biochem. Biophys. Res. Commun.* 76:836; Fraley (1979) *Proc. Natl. Acad. Sci. USA* 76:3348; Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76:145; Fraley (1980) *J. Biol. Chem.* (1980) 255:10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75:145; and Schaefer-Ridder (1982) *Science* 215:166.

E.Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; and Utermann (1984) *Hum Genet* 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phospholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol.* (*supra*); Pitas (1980) *J. Biochem.* 255:5454-5460 and Mahey (1979) *J Clin. Invest* 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Technologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

F. Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic acid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and putrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin[®], and lipofectAMINE[®] are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

Nucleic Acid Hybridisation

“Hybridization” refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [*supra*] Volume 2, chapter 9, pages 9.47 to 9.57.

“Stringency” refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200°C below the calculated T_m of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1 µg for a plasmid or phage digest to 10^{-9} to 10^{-8} g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy

yeast gene can be detected with an exposure time of only 1 hour starting with 1 μ g of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of 10^8 cpm/ μ g. For a single-copy mammalian gene a conservative approach would start with 10 μ g of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than 10^8 cpm/ μ g, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature (T_m) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$T_m = 81 + 16.6(\log_{10} C_i) + 0.4[\%(G + C)] - 0.6(\%\text{formamide}) - 600/n - 1.5(\%\text{mismatch}).$$

where C_i is the salt concentration (monovalent ions) and n is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

Nucleic Acid Probe Assays

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Probes may be produced by synthetic procedures, such as the triester method of Matteucci *et al.* [*J. Am. Chem. Soc.* (1981) 103:3185], or according to Urdea *et al.* [*Proc. Natl. Acad. Sci. USA* (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated *eg.* backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase *in vivo* half-life, alter RNA affinity, increase nuclease resistance *etc.* [*eg.* see Agrawal & Iyer (1995) *Curr Opin Biotechnol* 6:12-19; Agrawal (1996) *TIBTECH* 14:376-387]; analogues such as peptide nucleic acids may also be used [*eg.* see Corey (1997) *TIBTECH* 15:224-229; Buchardt *et al.* (1993) *TIBTECH* 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis *et al.* [*Meth. Enzymol.* (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al* [*supra*]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

EXAMPLES

The examples describe nucleic acid sequences which have been identified in *N. meningitidis*, and *N. gonorrhoeae* along with their respective and putative translation products. Not all of the nucleic acid sequences are complete *ie.* they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in *N. meningitidis*
- the putative translation product of said *N. meningitidis* sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from *N. gonorrhoeae*
- the putative translation product of said *N. gonorrhoeae* sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a corresponding nucleotide sequence identified from strain A of *N. meningitidis*
- the putative translation product of said *N. meningitidis* strain A sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (<http://www.ncbi.nlm.nih.gov>) using the algorithms BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, & tBLASTx [eg. see also Altschul *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed. Lower case letters represent ambiguities which arose during

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (<http://www.psort.nibb.ac.jp>). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (*e.g.* to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

Chromosomal DNA Preparation

N. meningitidis strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl₃/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, *Eco*RI-*Nde*I or *Eco*RI-*Nhe*I), depending on the restriction pattern of the gene of interest. The 3' primers included a *Xho*I or a *Hind*III restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using *Bam*HI-*Xho*I, *Bam*HI-*Hind*III, *Eco*RI-*Xho*I or *Eco*RI-*Hind*III), and pET21b+ (using *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III or *Nhe*I-*Hind*III).

5'-end primer tail:	<u>CGCGGATCCCATATG</u>	(<i>Bam</i> HI- <i>Nde</i> I)
	<u>CGCGGATCCGCTAGC</u>	(<i>Bam</i> HI- <i>Nhe</i> I)
	<u>CCGGAATTCTACATATG</u>	(<i>Eco</i> RI- <i>Nde</i> I)
	<u>CCGGAATTCTAGCTAGC</u>	(<i>Eco</i> RI- <i>Nhe</i> I)
3'-end primer tail:	<u>CCCGCTCGAG</u>	(<i>Xho</i> I)
	<u>CCCGCTCGAG</u>	(<i>Hind</i> III)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*Eco*RI, *Kpn*I or *Sal*I for the 5' primers and *Pst*I, *Xba*I, *Sph*I or *Sal*I for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) <u>AAAGAATTC</u>	(<i>Eco</i> RI)
	(AAA) <u>AAAGGTACC</u>	(<i>Kpn</i> I)
3'-end primer tail:	(AAA) <u>AAACTGCAG</u>	(<i>Pst</i> I)
	(AAA) <u>AAATCTAGA</u>	(<i>Xba</i> I)

$$T_m = 4 (G+C) + 2 (A+T) \quad (\text{tail excluded})$$

$$T_m = 64.9 + 0.41 (\% \text{ GC}) - 600/N \quad (\text{whole primer})$$

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in *Gonococcus* or in *Meningococcus A*. Hence, when the *Meningococcus B* sequence was incomplete or uncertain, *Gonococcal* or *Meningococcal A* sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40μM of each oligonucleotide primer, 400-800μM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl₂), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaq, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimised by the addition of 10µl DMSO or 50µl 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	Denaturation	Hybridisation	Elongation
First 5 cycles	30 seconds 95°C	30 seconds 50-55°C	30-60 seconds 72°C
Last 30 cycles	30 seconds 95°C	30 seconds 65-70°C	30-60 seconds 72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of 30µl or 50µl with either H₂O or 10mM Tris, pH 8.5.

Digestion of PCR fragments

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-

terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40 µl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30 µl or 50 µl with either H₂O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia). 10 µg plasmid was double-digested with 50 units of each restriction enzyme in 200 µl reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 µl of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD₂₆₀ of the sample, and adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

10 µg of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of 200 µl with the appropriate buffer overnight at 37°C. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD_{260nm} and the concentration adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

Cloning

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 µl, a molar

ratio of 3:1 fragment/vector was ligated using 0.5 μ l of NEB T4 DNA ligase (400 units/ μ l), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boehringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, 100 μ l *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800 μ l LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200 μ l of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100 μ g/ml ampicillin. The cells were then pelleted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30 μ l. 5 μ l of each individual miniprep (approximately 1g) were digested with either *NdeI/XhoI* or *BamHI/XhoI* and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20 μ l, that included 0.5 μ l T4 DNA ligase (400 units/ μ l, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boehringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100 μ l of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800 μ l LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200 μ l of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50µg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoRI-PstI* cloning sites, or *EcoRI-SalI*, or *SalI-PstI*. After cloning, the recombinant plasmids may be introduced in the *E.coli* host W3110.

Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 µl of each construct was used to transform 30 µl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 µg/ml) in 100 ml flasks, making sure that the OD₆₀₀ ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addition of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000g and the pellet resuspended in PBS for further use.

GST-fusion proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD₂₈₀ of 0.02-0.06. The GST-fusion protein was eluted by addition of 700µl cold Glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M'') (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD_{280nm} of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of 700µl cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD_{280nm} of the eluate indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

His-fusion soluble proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Ni²⁺-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the O.D₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

His-fusion insoluble proteins large-scale purification.

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to O.D₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150µl Ni²⁺-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD₂₈₀ of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

Purification of His-fusion proteins.

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni²⁺-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD_{280nm} of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the the OD_{280nm} of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

collected until the $O.D_{280nm}$ indicated all the recombinant protein was obtained. 20 μ l aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

His-fusion proteins renaturation

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200 μ g/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20 μ g/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

Purification of proteins

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500 μ l buffer M1 (PBS pH 7.2). 25 μ l of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M NaH₂ PO₄] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M NaH₂PO₄] overnight at 4°C. The

supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

Mice immunisations

20µg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with Al(OH)₃ as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20µg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

ELISA assay (sera analysis)

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN₃ in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 µl of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 µl of H₂O) were added to each well and the plates were left at room temperature for 20 minutes. 100 µl H₂SO₄ was added to each well and OD₄₉₀ was followed. The ELISA was considered positive when OD₄₉₀ was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN₃ in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-

phenildiamine and 10 μ l of H₂O₂) were added to each well and the plates were left at room temperature for 20 minutes. 100 μ l of 12.5% H₂SO₄ was added to each well and OD₄₉₀ was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD₄₉₀ value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD₄₉₀ of 0.4 was higher than 1:400.

FACScan bacteria Binding Assay procedure.

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN₃) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD₆₂₀ of 0.07. 100 μ l bacterial cells were added to each well of a Costar 96 well plate. 100 μ l of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200 μ l/well of blocking buffer in each well. 100 μ l of R-Phicoerytrin conjugated F(ab)₂ goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200 μ l/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200 μ l/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

OMV preparations

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice (50% duty cycle, 50% output). Unbroken cells were removed by centrifugation at 5000g for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000g at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000g for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

Whole Extracts preparation

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

Western blotting

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

Bactericidal assay

MC58 and 2996 strains were grown overnight at 37°C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37°C on a nutator and let to grow until OD₆₂₀ was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD₆₂₀ of 0.5, diluted 1:20000 in Gey's buffer and stored at 25°C.

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various *Neisseria* strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1	(forward) CGAATCCGGACGGCAGGACTC
orf 4.3	(reverse) GGCAGGGAATGGCGGATTAAAG
919.1	(forward) AAAATGCCTCTCCACGGCTG or CTGCGCCCTGTGTTAAAATCCCCT
919.6	(reverse) CAAATAAGAAAGGAATTTTG or GGTATCGCAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at ~ 54° or ~ 60° (in according to Tm of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

orf 4.1 (forward)	CGAATCCGGACGGCAGGACTC
orf 4.2 (forward)	CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse)	GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse)	TCTTTGAGTTTGATCCAACC

- 919.1 (forward) AAAATGCCTCTCCACGGCTG or
CTGCGCCCTGTGTTAAAATCCCCT
- 919.2 (forward) ATCCTTCCGCCTCGGCTGCG
- 919.3 (forward) AAAACAGCGGCACAATCGAC
- 919.4 (forward) ATAAGGGCTACCTCAAATC
- 919.5 (forward) GCGCGTGGATTATTTTGGG
- 919.6 (reverse) CAAATAAGAAAGGAATTTTG or
GGTATCGCAAACTTCGCCTTAATGCG
- 919.7 (reverse) CCCAAGGTAATGTAGTGCCG
- 919.8 (reverse) TAAAAAAAAGTTCGACAGGG
- 919.9 (reverse) CCGTCCGCCTGTCGTCGCCC
- 919.10 (reverse) TCGTTCCGGCGGGGTCGGGG

All documents cited herein are incorporated by reference in their entireties.

The following Examples are presented to illustrate, not limit, the invention

EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 1: Oligonucleotides used for PCR for Examples 2-10

ORF	Primer	Sequence	Restriction sites
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT <SEQ ID 3021>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <SEQ ID 3022>	XhoI
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA <SEQ ID 3023>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <SEQ ID 3024>	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT <SEQ ID 3025>	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC <SEQ ID 3026>	XhoI
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC <SEQ ID 3027>	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <SEQ ID 3028>	XhoI
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT <SEQ ID 3029>	BamHI-NdeI

128	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <SEQ ID 3030>	XhoI
	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <SEQ ID 3031>	BamHI-NdeI
206	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA <SEQ ID 3032>	XhoI
	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <SEQ ID 3033>	BamHI-NdeI
287	Reverse	CCCGCTCGAG-TTCTGTAAAAAAGTATGTGC <SEQ ID 3034>	XhoI
	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <SEQ ID 3035>	EcoRI-NheI
406	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC <SEQ ID 3036>	XhoI
	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG <SEQ ID 3037>	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGCTCTATG <SEQ ID 3038>	XhoI

Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from *N. gonorrhoeae*, "m" means a sequence from *N. meningitidis B*, and "a" means a sequence from *N. meningitidis A*; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an *N. gonorrhoeae* DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unsuffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a *N. gonorrhoeae* sequence or a *N. meningitidis A* sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

72

designated sequences being compared, the designated sequence controls and is the actual sequence being compared.

ORF: contig:

279 gnm4.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3039>:

m279.seq

1	ATAACGCGGA	TTTGCGGCTG	CTTGATTTC	ACGGTTTTCA	GGGCTTCGGC
51	AAGTTTGTCT	GCGGCGGGTT	TCATCAGGCT	GCAATGGGAA	GGTACGGACA
101	CGGGCAGCGG	CAGGGCGCGT	TTGGCACCGG	CTTCTTTGGC	GGCAGCCATG
151	GCGCGTCCGA	CGGCGGCGGC	GTTGCCTGCA	ATCACGATTT	GTCCGGGTGA
201	GTTGAAGTTG	ACGGCTTCGA	CCACTTCGCT	TTGGGCGGCT	TCGGCACAAA
251	TGGCTTTAAC	CTGCTCATCT	TCCAAGCCGA	GAATCGCCGC	CATTGCGCCG
301	ACGCCTTGCG	GTACGGCGGA	CTGCATCAGT	TCGGCGCGCA	GGCGCACAGG
351	TTTGACCGCG	TCGGCAAAAT	TCAATGCGCC	GGCGGCAACG	AGTGCGGTGT
401	ATTGCGCGAG	GCTGTGTCCG	GCAACGCGCG	CAGGCGTTTT	GCCGCCCGCT
451	TCTAAATAG				

This corresponds to the amino acid sequence <SEQ ID 3040; ORF 279>:

m279.pep

```

1  ITRICGCLIS TVFRASASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM
51  ARPTAAALPA ITICPGELKL TASTTSLWAA SAQMALTCSS SKPRIAAIAP
101 TPCGTADCIS SARRRSLTA SAKFNAPAAT SAVYSPRLCP ATAAGVLP
151 SK*

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3041>:

g279.seq

1	atgacgcgga	tttgcggtcg	cttgatttca	acggttttga	gtgtttcggc
51	aagtttgctg	gcggcggtt	tcatacaggct	gcaatgggaa	ggaacggata
101	ccggcagcgg	cagggcgctg	ttggctccgg	cttctttggc	ggcagccatg
151	gtgctgccga	cggcgcgcg	gttgctcgca	atcacgactt	gtccgggcga
201	gttgaagtbg	acggcttcga	ccacttcgcc	ctgtgcggat	tcggracaaa
251	tctgcctgac	ctgttcattc	tccaaaccca	aaatggcgc	cattgcgcct
301	acgccttgcg	gtacggcgga	ctgcatcagt	tcggcgcgca	ggcggacgag
351	tttcagcgga	tcggcaaaat	ccaatgcctc	ggcggcgaca	agcgcggtgt
401	attgcgcgag	gctgtgtccg	gcaacggcgg	caggcgtttt	gccgccact
451	tccaaataq				

This corresponds to the amino acid sequence <SEQ ID 3042; ORF 279.ng>:

g279.pep

```

1  MTRICGCLIS TVLSVSASLS AAGFIRLQWE GTDTGSGRRAR LAPASLAAAM
51  VRPTAAALPA ITTCPGELKL TASTTSPCAD SAQICLTCSS SKPKMAAIAP
101 TPCGTADCIS SARRRTSLTA SAKSNASAAT SAVYSPRLCP ATAAGVLPPT
151 SK*

```

ORF 279 shows 89.5% identity over a 152 aa overlap with a predicted ORF (ORF 279.ng) from *N. gonorrhoeae*:

```

      10          20          30          40          50          60
m279.pep  ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMARPTAAALPA
           :|:|||||:|:|||||||:|||||||:|||||||:|||||||:|||||||:
g279       MTRICGCLISTVLVSASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMVRPTAAALPA
           10          20          30          40          50          60

```

	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA					
g279	ITTCPGELKLTASTTSPCADSAQICLTCSSSKPKMAAIAPTPCGTADCISSARRRTSLTA					
	70	80	90	100	110	120
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLP PASKX					
g279	SAKSNASAATSAVYSPRLCPATAAGVLP PPTSKX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3043>:

```
a279.seq
1  ATGACNCNGA  TTTGCGGCTG  CTTGATTTC ACGGTTT NNA GGGCTTCGGC
51  GAGTTTGTGCG GCGGCGGGTT TCATGAGGCT GCAATGGGAA GGTACNGACA
101  CNGGCAGCGG  CAGGGCGCGT TTGGCGCGCG CTTCTTTGCG GGCAAGCATA
151  GCGCGCTCGA  CGGCGGCGGC ATTGCCTGCA ATCACGACTT GTCCGGGCGA
201  GTTGAAGTTG  ACGGCTTCAA CCACTTCATC CTGTGCGGAT TCGGCGCAAA
251  TTTGTTTTAC  CTGTTTCATCT TCCAAGCCGA GAATCGCCGC CATTGCGCCC
301  ACGCCTTGCG  GTACGGCGGA CTGCATCAGT TCGGCGCGCA NGCGCACGAG
351  TTTGACCGCG  TCGGCAAAAT CCAATGCGCC GCGGCAACN AGTGCGGTGT
401  ATTCGCCGAN  GCTGTGTCCG GCAACGGCGG CAGGCGTTT GCCGCCCGCT
451  TCCGAATAG
```

This corresponds to the amino acid sequence <SEQ ID 3044; ORF 279.a>:

```
a279.pep
1  MTXICGCLIS TVXRASASLS AAGFMRLQWE GTDTGSGRAR LAPASLAASI
51  ARSTAAALPA ITTCPGELKL TASTTSSCAD SAQICFTCSS SKPRIAAIAP
101  TPCGTADCIS SARXRTSLTA SAKSNAPAAT SAVYSPXLCP ATAAGVLPPA
151  SE*
```

m279/a279 ORFs 279 and 279.a showed a 88.2% identity in 152 aa overlap

	10	20	30	40	50	60
m279.pep	ITRICGLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAMARPTAAALPA					
a279	MTXICGCLISTVXRASASLSAAGFMRLQWEGTDTGSGRARLAPASLAASIARSTAAALPA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA					
a279	ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARXRTSLTA					
	70	80	90	100	110	120
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLP PASKX					
a279	SAKSNAPAATSAVYSPXLCPATAAGVLP PASEX					
	130	140	150			

519 and 519-1 gnm7.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3045>:

```
m519.seq (partial)
1  ..TCCGTTATCG GCGTATGGA GTTGGACAAA ACGTTTGAAG AACGCGACGA
51  AATCAACAGT ACTGTTGTTG CGGCTTTGGA CGAGGCGGCC GGGgCTTgGG
101  GTGTGAAGGT TTTGCGTTAT GAGATTAAAG ACTTGTTCC GCCGCAAGAA
151  ATCCTTCGCT CAATGCAGGC GCAAATTACT GCCGAACGCG AAAAACGCGC
201  CCGTATCGCC GAATCCGAAG GTCGTAAAAT CGAACAAATC AACCTTGCCA
```



```

251 GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
301 GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
351 AGGTGAAGCG GAATCCTTGC GCCTTGTTCG CGAAGCCAAT GCCGAAGCCA
401 TCCGTCAAAT TGCCGCCGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
451 AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
501 AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTTGCCGAC ATCGGCAGCC
551 TGATTTCTGC CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA

```

This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:

```

m519.pep (partial)
1  .SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
51  ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLAGQREAE IQQSEGEAQA
101 AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEAIRQIAAA LQTQGGADAV
151 NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSLISAGMK IIDSSKTAK*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3047>:

```

g519.seq
1  atggaatddd tcattatcctt gttggcagcc gtcgccgttt tcggcttcaa
51  atcctttgtc gtcattcccc agcaggaagt ccacgtttgtc gaaaggctcg
101 ggcgtttcca tcgcgccttg acggccggtt tgaatatddd gattcccttt
151 atcgaccgag tcgcctaccg ccattcgctg aaagaaatcc ctttagacgt
201 acccagccag gtctgcatca cgcgcgataa tacgcaattg actgttgacg
251 gcatcatcta tttccaagta accgatccca aactcgccctc atacggttcg
301 agcaactaca ttatggcaat taccagctt gcccaaacga cgctgcgttc
351 cgttatcggg cgtatggagt tggacaaaac gtttgaagaa cgcgacgaaa
401 tcaacagtac cgtcgtctcc gccctcgatg aagccgccgg ggcttggggg
451 gtgaaagtcc tccgttacga aatcaaggat ttggttcgcg cgcaagaaat
501 ccttcgcgca atgcaggcac aaattaccgc cgaacgcgaa aaacgcgccc
551 gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ccttgccagt
601 ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctccaggtgc
651 ggtcaatgag tccaatgccg agaaaatcgc ccgcatcaac cgcgccaag
701 gcgaagcgga atccctgcgc cttgttgccg aagccaatgc cgaagccaac
751 cgtcaaatgg ccgcccgcct tcaaacccaa agcggggcgg atgcgggtcaa
801 tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaag
851 aagacaatac gcggattaag cccgccaagg ttgccgaaat cgggaaccct
901 aattdtcggc ggcataaaaa attdtcgcca gaagcaaaaa cggccaataa
951 a

```

This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:

```

g519.pep
1  MEFFIILLAA VAVFGFKSFV VIPQOEHVHV ERLGRFHRAL TAGLNILIPF
51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEEAQAANA SNAEKIARIN RAKGEAESLR LVAEANAAN
251 RQIAAALQTQ SGADAVNLKI AGQYVTAFKN LAKEDNTRIK PAKVAEIGNP
301 NFRRHEKFSP EAKTAK*

```

ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng) from *N. gonorrhoeae*:

```

m519/g519

m519.pep                               10      20      30
SVIGRMELDKTFEERDEINSTVVAALDEAA
|||||
g519      YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
          90      100     110     120     130     140

          40      50      60      70      80      90
m519.pep      GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLAGQREAE

```

```
a519.seq
1  ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGGTTCAA
51  ATCCTTTGTT GTCATCCCAAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GGC GTTTCCTA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTC
301 AGCAACTACA TTAGGCGGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTG
351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
451 GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGTG CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAAATGCG TCAAAATGCC AGAAAATCGC CGGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GCGGTTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCCCG GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCGCCAATG TTGCCGACAT CGGCACGCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA
```

```
a519.pep
1  MEFFFIILLAA VVVFGEKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRISVIG RMELDKTFEE REINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAAQAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLM PANVADIGSL
301 ISAGMKIIDS SKTAK*
```

```

                                     10      20      30
m519.pep                          SVIGRMELDKTFEERDEINSTVVAALDEAA
                                     |||:|||:|||:|||:|||:|||:|||:|||:|||
a519                               YFQVTDPKLASYGSSNYIMAITQLAQTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
                                   90      100      110      120      130      140

                                   40      50      60      70      80      90
m519.pep                          GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                                   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||
a519                              GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                                   150      160      170      180      190      200

                                   100      110      120      130      140      150

```

```

m519.pep      IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
a519          IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
                210      220      230      240      250      260

m519.pep      NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
a519          NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
                270      280      290      300      310

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3051>:

```

m519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCCGTTT TCGGTTTCAA
51 ATCCTTTGTT GTCATCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
101 GCGTTTCCA TCGCGCCCTG ACGGCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC TGTTGTTGCG GCTTTGGACG AGGCGGCCGG GGCTTGGGGT
451 GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC CGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCGCCCTT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAT TATCGACAGC AGCAAACCG CCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```

m519-1.
1  MEFFIILLVA VAVFGFKSFV VIPQOEHVHV ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSI KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 QQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3053>:

```

g519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGCAGCC GTCGCCGTTT TCGGCTTCAA
51 ATCCTTTGTC GTCATCCCCC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGTTTCCA TCGCGCCCTG ACGGCCGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGATAA TACGCAATTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGATCCCA AACTCGCCTC ATACGGTTCG
301 AGCAACTACA TTATGGCAAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC CGTCGTCTCC GCCCTCGATG AAGCGGCCGG GGCTTGGGGT
451 GTGAAAGTCC TCCGTTACGA AATCAAGGAT TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCGCA ATGCAGGCAC AAATTACCGC CGAACGCGAA AAACGCGCCC
551 GTATTGCCGA ATCCGAAGGC CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGTG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCCAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GCGAAGCGGA ATCCCTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCGCCCTT TCAAACCCAA GGCGGGGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTAGCCGC GTTCAACAAT CTTGCCAAAG

```

851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GCATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>:

g519-1.pep
1 MEFFIILLAA VAVFGFKSFV VIPQQEVHV V ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKCLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQA AVNA SNAEKIARIN RAKGEAESLR LVAEANA EAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

m519-1/g519-1 ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap

	10	20	30	40	50	60
g519-1.pep	MEFFIILLAAVAVFGFKSFV	VIPQQEVHVVERLGRFHRAL	TAGLNILIPFIDRVAYRHSL			
m519-1	MEFFIILLVAVAVFGFKSFV	VIPQQEVHVVERLGRFHRAL	TAGLNILIPFIDRVAYRHSL			
	10	20	30	40	50	60
g519-1.pep	KEIPLDVPSQVCITRDNTQL	TVDGIIFYQVTDPKCLASYGS	SNYIMAITQLAQTTLRSVIG			
m519-1	KEIPLDVPSQVCITRDNTQL	TVDGIIFYQVTDPKCLASYGS	SNYIMAITQLAQTTLRSVIG			
	70	80	90	100	110	120
g519-1.pep	KEIPLDVPSQVCITRDNTQL	TVDGIIFYQVTDPKCLASYGS	SNYIMAITQLAQTTLRSVIG			
m519-1	KEIPLDVPSQVCITRDNTQL	TVDGIIFYQVTDPKCLASYGS	SNYIMAITQLAQTTLRSVIG			
	70	80	90	100	110	120
g519-1.pep	RMELDKTFEERDEINSTVVS	ALDEAAGAWGVKVLRYEIKD	LVPPQEILRAMQQAQITAERE			
m519-1	RMELDKTFEERDEINSTVVA	ALDEAAGAWGVKVLRYEIKD	LVPPQEILRSMQQAQITAERE			
	130	140	150	160	170	180
g519-1.pep	RMELDKTFEERDEINSTVVS	ALDEAAGAWGVKVLRYEIKD	LVPPQEILRAMQQAQITAERE			
m519-1	RMELDKTFEERDEINSTVVA	ALDEAAGAWGVKVLRYEIKD	LVPPQEILRSMQQAQITAERE			
	130	140	150	160	170	180
g519-1.pep	KRARIAESEGRKIEQINLAS	GQREAEIQQSEGEAQA AVNA	SNAEKIARINRAKGEAESLR			
m519-1	KRARIAESEGRKIEQINLAS	GQREAEIQQSEGEAQA AVNA	SNAEKIARINRAKGEAESLR			
	190	200	210	220	230	240
g519-1.pep	KRARIAESEGRKIEQINLAS	GQREAEIQQSEGEAQA AVNA	SNAEKIARINRAKGEAESLR			
m519-1	KRARIAESEGRKIEQINLAS	GQREAEIQQSEGEAQA AVNA	SNAEKIARINRAKGEAESLR			
	190	200	210	220	230	240
g519-1.pep	LVAEANA EAI RQIAAALQTQ	GGADAVNLKIAEQYVAAFNN	LAKESNTLIMPANVADIGSL			
m519-1	LVAEANA EAI RQIAAALQTQ	GGADAVNLKIAEQYVAAFNN	LAKESNTLIMPANVADIGSL			
	250	260	270	280	290	300
g519-1.pep	LVAEANA EAI RQIAAALQTQ	GGADAVNLKIAEQYVAAFNN	LAKESNTLIMPANVADIGSL			
m519-1	LVAEANA EAI RQIAAALQTQ	GGADAVNLKIAEQYVAAFNN	LAKESNTLIMPANVADIGSL			
	250	260	270	280	290	300
g519-1.pep	ISAGMKIIDSSKTAKX					
m519-1	ISAGMKIIDSSKTAKX					
	310					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3055>:

a519-1.seq
1 ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA
51 ATCCTTTGTT GTCATCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGCTTTCCA TCGCGCCCTG ACGGCCGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGGACAA TACGCAGCTG ACTGTTGACG
251 GTATCATCTA TTTCCAAGTA ACCGACCCA AACTCGCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA

```

401 TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
451 GTGAAGGTTT TGC GTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTACGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GCGCGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3056; ORF 519-1.a>:

a519-1.pep.

```

1  MEFFIILLAA VVVFQFVSVV VIPOQEVHVV ERLGRFHRAL TAGLNILIPF
51  IDRVAIRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAELI
251 RQIAAALQTO GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

m519-1/a519-1 ORFs 519-1 and 519-1.a showed a 99.0% identity in 315 aa overlap

	10	20	30	40	50	60
a519-1.pep	MEFFIILLAAVVVFQFVSVVVIPOQEVHVVVERLGRFHRALTAGLNILIPFIDRVAIRHSL					
m519-1	MEFFIILLVAVAVFQFVSVVVIPOQEVHVVVERLGRFHRALTAGLNILIPFIDRVAIRHSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
a519-1.pep	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
m519-1	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
	70	80	90	100	110	120
	130	140	150	160	170	180
a519-1.pep	RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE					
m519-1	RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE					
	130	140	150	160	170	180
	190	200	210	220	230	240
a519-1.pep	KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR					
m519-1	KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR					
	190	200	210	220	230	240
	250	260	270	280	290	300
a519-1.pep	LVAEANAELIRQIAAALQTOGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
m519-1	LVAEANAELIRQIAAALQTOGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
	250	260	270	280	290	300
	310					
a519-1.pep	ISAGMKIIDSSKTAKX					
m519-1	ISAGMKIIDSSKTAKX					
	310					

576 and 576-1 gnm22.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3057>:

m576.seq.. (partial)

```

1   ..ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
51  GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
101 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
151 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
201 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
251 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
301 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
351 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
401 TCGACAGCAG CAAAGCCAAC GCGGCCCGG TCACCTTCCC TTTGAGCCAA
451 GTGATTCCGG GTTGGACCGA AGCGTACAG CTTCTGAAAG AAGGCGGCGA
501 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTCCGG
551 GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
601 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
651 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3058; ORF 576>:

m576.pep.. (partial)

```

1   ..MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
51  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
101 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
201 KIGAPENAPA KQPAQVDIKK VN*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3059>:

g576.seq.. (partial)

```

1   ..atgggcggtg acatcggacg ctccctgaaa caaatgaagg aacagggcgc
51  ggaaatcgat ttgaaagtct ttaccgatgc catgcaggca gtgtatgacg
101 gcaaagaaat caaaatgacc gaagagcagg cccaggaaat gatgatgaaa
151 ttcctgcagg agcagcaggc taaagccgta gaaaaacaca aggcggatgc
201 gaaggccaac aaagaaaaag gcgaagcctt cctgaaggaa aatgccgcgc
251 aagacggcgt gaagaccact gcttccggtc tgcagtacaa aatcaccaaa
301 cagggtgaag gcaaacagcc gacaaaagac gacatcgta ccgtggaata
351 cgaaggccgc ctgattgacg gtaccgtatt cgacagcagc aaagccaacg
401 ggcggccgcg caccttcctt ttgagccaag tgattccggg ttggaccgaa
451 ggctacggc ttctgaaaga aggcggcgaa gccacgttct acatcccgtc
501 caaccttgcc taccgcgaac aggggtcggg cgaaaaaatc ggtccgaacg
551 ccactttggt atttgacgtg aaactggtca aaatcggcgc acccgaaaac
601 gcgcccgcga agcagccgga tcaagtcgac atcaaaaaag taaattaa

```

This corresponds to the amino acid sequence <SEQ ID 3060; ORF 576.ng>:

g576.pep.. (partial)

```

1   ..MGVDIGRSLK QMKEQGAEID LKVFTDAMQA VYDGKEIKMT EEQAQEVMMK
51  FLQEQQAKAV EKHKADAKAN KEKGEAFLKE NAAEDGVKTT ASGLQYKITK
101 QGEGKQPTKD DIVTVEYEGR LIDGTVFDSS KANGGPATFP LSQVIPGWTE
151 GVRLLKEGGE ATFYIPSNLA YREQGAGEKI GPNATLVFDV KLVKIGAPEN
201 APAKQPDQVD IKKVN*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m576/g576 ORFs 576 and 576.ng showed a 97.2% identity in 215 aa overlap

```

              10      20      30      40      50      60
m576.pep      MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQAQEVMMKFLQ
              |||:|||||:|||||:|||||:|||||:|||||

```

80

```

g576          MGVDIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQ
                10      20      30      40      50

                70      80      90      100     110     120
m576.pep      EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQEGEKQPTKDDIV
                |||
g576          EQQAKAVEKHKADAKANKEKGEAFLKENAAEDGVKTTASGLQYKITKQEGEKQPTKDDIV
                60      70      80      90      100     110

                130     140     150     160     170     180
m576.pep      TVEYEGRLIDGTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYRE
                |||
g576          TVEYEGRLIDGTVFDSSKANGGPATFPLSQVIPGWTEGVRLLEKGEATFYIPSNLAYRE
                120     130     140     150     160     170

                190     200     210     220
m576.pep      QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN
                |||
g576          QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN
                180     190     200     210

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3061>:

```

a576.seq
1   ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCGG CCGCTTTGGC
51  ACTTTCGCC  TCGGGCAAAA AAGAAGCCGC CCGCGCATCT GCATCCGAAC
101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC  CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3062; ORF 576.a>:

```

a576.pep
1   MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
151 LQYKITQGE  GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

m576/a576 ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap

```

                10      20      30
m576.pep      MQQASYAMGVDIGRSLKQMKEQGAEIDLKV
                |||
a576          CGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGVDIGRSLKQMKEQGAEIDLKV
                30      40      50      60      70      80

                40      50      60      70      80      90
m576.pep      FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                |||
a576          FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                90      100     110     120     130     140

```

	100	110	120	130	140	150
m576.pep	KDGVKTTASGLQYKITKQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
a576	KDGVKTTASGLQYKITKQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
	150	160	170	180	190	200
	160	170	180	190	200	210
m576.pep	VIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA					
a576	VILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA					
	210	220	230	240	250	260
	220					
m576.pep	KQPAQVDIKKVN					
a576	KQPAQVDIKKVN					
	270					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3063>:

```

m576-1.seq
1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCGG CCGCTTTGGC
51  ACTTTCGGCC TGCGGCAAAA AAGAAGCCGC CCCC GCATCCGAAC
101 CTGCCGCCGC TTCTTCGGCG CAGGGCGACA CCTCTTCGAT CGGCAGCAGC
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCCGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

```

m576-1.pep
1  MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
151 LQYKITQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3065>:

```

g576-1.seq
1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCGG CCGCTTTGGC
51  ACTTTCGGCC TGCGGCAAAA AAGAAGCCGC CCCC GCATCCGAAC
101 CTGCCGCCGC TTCTGCCGCG CAGGGCGACA CCTCTTCAAT CGGCAGCAGC
151 ATGCAGCAGG CAAGCTATGC AATGGGCGTG GACATCGGAC GCTCCCTGAA
201 ACAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGATG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCCCAGGAAG TGATGATGAA ATTCCTGCAG GAGCAGCAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGATG CGAAGGCCAA CAAAGAAAAA GGCGAAGCCT
401 TCCTGAAGGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGT
451 CTGCAGTACA AAATCACCAA ACAGGGTGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACCGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG CCACCTTCCC TTTGAGCCAA
601 GTGATTCCGG GTTGGACCGA AGGCGTACGG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG

```


701 GCGAAAAAAT CCGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG ATCAAGTCGA
 801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>:

g576-1.pep

1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASAA QGDTSSIGST
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ
 101 AQEVMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
 151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ
 201 VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVDVKLV
 251 KIGAPENAPA KQPDQVDIKK VN*

g576-1/m576-1 ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa overlap

	10	20	30	40	50	60
g576-1.pep	MNTIFKISAL	TL	SAALALSA	CGKKEAAPAS	ASEPAASAA	QGDTSSIGSTMQQASYAMGV
m576-1	MNTIFKISAL	TL	SAALALSA	CGKKEAAPAS	ASEPAASSA	QGDTSSIGSTMQQASYAMGV
	10	20	30	40	50	60
	70	80	90	100	110	120
g576-1.pep	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEEQ	AQEVMKFLQ	EQQAKAVEKH
m576-1	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEEQ	AQEVMKFLQ	EQQAKAVEKH
	70	80	90	100	110	120
	130	140	150	160	170	180
g576-1.pep	KADAKANKEK	GEAFLENAA	KDGVKTTASG	LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
m576-1	KADAKANKEK	GEAFLENAA	KDGVKTTASG	LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
	130	140	150	160	170	180
	190	200	210	220	230	240
g576-1.pep	GTVFDSSKAN	GGPATFPLSQ	VIPGWTEGVR	LLKEGGEATF	YIPSNLAYRE	QGAGEKIGPN
m576-1	GTVFDSSKAN	GGPATFPLSQ	VIPGWTEGVR	LLKEGGEATF	YIPSNLAYRE	QGAGDKIGPN
	190	200	210	220	230	240
	250	260	270			
g576-1.pep	ATLVFDVKLV	KIGAPENAPA	KQPDQVDIKK	VNX		
m576-1	ATLVFDVKLV	KIGAPENAPA	KQPAQVDIKK	VNX		
	250	260	270			

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3067>:

a576-1.seq

1 ATGAACACCA TTTTCAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
 51 ACTTTCCGCC TCGGCAAAA AAGAAGCCGC CCCCAGCATCT GCATCCGAAC
 101 CTGCCCGCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
 151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
 201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
 251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
 301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
 351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
 401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
 451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
 501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
 551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTAGCCAA
 601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
 651 AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG

701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
 801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:

a576-1.pep

1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
 101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
 151 LQYKITQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
 201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
 251 KIGAPENAPA KQPAQVDIKK VN*

a576-1/m576-1 ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa overlap

a576-1.pep	10	20	30	40	50	60
	MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV					
m576-1	10	20	30	40	50	60
	MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV					
a576-1.pep	70	80	90	100	110	120
	DIGRSLKQMKEQGAEIDLKVTEAMQAVYDGKEIKMTEEQAEVMMKFLQEQQAKAVEKH					
m576-1	70	80	90	100	110	120
	DIGRSLKQMKEQGAEIDLKVTEAMQAVYDGKEIKMTEEQAEVMMKFLQEQQAKAVEKH					
a576-1.pep	130	140	150	160	170	180
	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQEGGKQPTKDDIVTVEYEGRLID					
m576-1	130	140	150	160	170	180
	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQEGGKQPTKDDIVTVEYEGRLID					
a576-1.pep	190	200	210	220	230	240
	GTVFDSSKANGGPVTFPLSQVILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN					
m576-1	190	200	210	220	230	240
	GTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN					
a576-1.pep	250	260	270			
	ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN					
m576-1	250	260	270			
	ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN					

919 gnm43.seq

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3069>:

m919.seq

1 ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TACGGCATCG CCGCCGCCAT
 51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
 101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
 151 GGAACGACGG TCGGCGGCGG CGGGGCCGTC TATACCGTTG TACCGCACCT
 201 GTCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
 251 TCCGCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
 301 TGCGCCCAAG CCTTTCAAAC CCCCCTCCAT TCCTTTCAGG CAAAACAGTT
 351 TTTTGAACGC TATTTACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
 401 CCGGTACGGT TACCGGTAT TACGAACCGG TGCTGAAGGG CGACGACAGG

```

451 CGGACGGCAC AAGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACCGCCG ACCTCTCCcG ATTCCCCATC ACCGCGCGCA CAACAGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCGATACT CGGTACGCC
751 GAAGACCCCTG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCyTA CGTTTCCATC GGACGCTATA TGGCGGATAA GGGCTACCTC
901 AAACCTCGAC AAACCTCCAT GCAGGGCATT AAGTCTTATA TGCGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051 ACGCCGCTGA TGGGGGAATA TGCCGGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGTGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTC CTACCCAACG
1301 GTATGAAGCC CGAATACCGC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

```

m919 . pep
  1 MKKYLFRAL YGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
 51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGRSFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQSGSRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KSYMQRNPQR LAEVLGQNPS YIFFRELAYS SNDGPVGA LG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3071>:

```

g919 . seq
  1 ATGAAAAAAC ACCTGCTCCG CTCCGCCCTG TACGGcatCG CCGCCgccAT
 51 CctcgCCGCC TGCCAAAgca gGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG CCGGCATCCC CGACCCCGCC
151 GGAACGACGG TTGCCGGCGG CGGGGCCGTC TATACCGTTG TGCCGCACCT
201 GTCCATGCCC CACTGGGCGG CGCaggATTT TGCCAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TGCGCCCAAG CCTTTCAAAC CCCCGTGCAT TCCTTTCAGG CAAAGcGgTT
351 TTTTGAACGC TATTTACGC cgtGGCaggt tgcaggcaAC GGAAGcCTTG
401 Caggtacggt TACCGGTAT TACGAACCGG TGCTGAAGGG CGACGGCAGG
451 CGGACGGAAC GGGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTGCGGGG CGGAAAAAAC CTTGTCCGCA
551 TCAGGCAGac ggGAAAAAAC AGCGGCACGA TCGACAATGC CGGCGGCACG
601 CATACCGCCG ACCTCTCCCG ATTCCCCATC ACCGCGCGCA CAACGGcaat
651 caaaGGCAGG TTTGAaggAA GCCGCTTCCT CCCTTACCAC ACGCGCAACC
701 AAAtcaacGG CGGCgcgcTT GACGGCAAag cccCATCCT CggttacgcC
751 AagaccCcG tcgaacttTT TTTCATGCAC AtccaaggCT CGGGCCGCTT
801 GAAAACCCcg tccggcaaat acatCCGCAt cggaTacgcc gacAAAAACG
851 AACAtccgTa tgtttccatc ggACGctaTA TGGCGGACAA AGGCTACCTC
901 AAGctcgggc agACCTCGAT GCAGGgcac aaagcCTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC GGCAATGAGG GCCCGTCGG CGCACTGGGC
1051 ACGCCACTGA TGGGGGAATA CGCCGGCGCA ATCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CAGGCAGCGC GATCAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGTTAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATACGT CTGGCAGCTC CTGCCCCAACG
1301 GCATGAAGCC CGAATACCGC CCGTGA

```

ORF 919 shows 95.9 % identity over a 441 aa overlap with a predicted ORF (ORF 919.ng) from *N. gonorrhoeae*:

	10	20	30	40	50	60
m919.pep	MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV					
	: : : : : : : : : : : : : : : : : :					
g919	MKKHLLRSALYGIAAAILAACQSRSIQTFPPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m919.pep	YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER					
	: : : : :					
g919	YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER					
	70	80	90	100	110	120
	130	140	150	160	170	180
m919.pep	YFTPWQVAGNGSLAGTVTGYEYPVLKGDGRRRTAQARFPPIYGIPDDFISVPLPAGLRSGKA					
	: : : : :					
g919	YFTPWQVAGNGSLAGTVTGYEYPVLKGDGRRTERRARFPPIYGIPDDFISVPLPAGLRGGKN					
	130	140	150	160	170	180
	190	200	210	220	230	240
m919.pep	LVRIRQTGKNSGTIDNTGGTHADLSRFPI TARTTAIKGRFEGS RFLPYHTRNQINGGAL					
	: : : : :					
g919	LVRIRQTGKNSGTIDNAGGTHADLSRFPI TARTTAIKGRFEGS RFLPYHTRNQINGGAL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m919.pep	DGKAPILGYAEDPVELFFMHIIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL					
	: : : : :					
g919	DGKAPILGYAEDPVELFFMHIIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL					
	250	260	270	280	290	300
	310	320	330	340	350	360
m919.pep	KLGQTSMQGIKS YMRQN PQR LAEV LGQNPSYIFFREL AGSSNDGPVGALGTPLMGEYAGA					
	: : : : :					
g919	KLGQTSMQGI KAYMRQN PQR LAEV LGQNPSYIFFREL AGSGNEGVPVGALGTPLMGEYAGA					
	310	320	330	340	350	360
	370	380	390	400	410	420
m919.pep	VDRHYITLGAPLFVATAHPVTRKALNRLIM AQD TGSAI KGAVRV D YFWGYGDEAGELAGK					
	: : : : : :					
g919	IDRHYITLGAPLFVATAHPVTRKALNRLIM AQD TGSAI KGAVRV D YFWGYGDEAGELAGK					
	370	380	390	400	410	420

```

                430      440
m919.pep      QKTTGYVWQLLPNGMKPEYRPX
                |||||
g919          QKTTGYVWQLLPNGMKPEYRPX
                430      440

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3073>:

```

a919.seq
1  ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TGC GG CATCG CCGCCGCCAT
51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
151 GGAACGACGG TCGGCGGCGG CGGGGCCGTT TATACCGTTG TGCCGCACCT
201 GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TCGGCCAAG CCTTCAAAC CCCGTCCAT TCCGTT CAGG CAAAACAGTT
351 TTTGAACGC TATTTCACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
401 CCGGTACGGT TACCGGCTAT TACGAGCCGG TGCTGAAGGG CGACGACAGG
451 CGGACGGCAC AAGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATAACGCCG ACCTCTCCCA ATCCCCATC ACTGCGCGCA CAACGGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCATACT CGGTTACGCC
751 GAAGACCCCG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAACG
851 AACATCCCTA CGTTTCCATC GGACGCTATA TGGCGGACAA AGGCTACCTC
901 AAGCTCGGGC AGACCTCGAT GCAGGCATC AAAGCCTATA TGCAGCAAAA
951 CCCGCAACGC CTCGCCGAAG TTTTGGGGCA AAACCCAGC TATATCTTTT
1001 TCCGAGAGCT TACCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051 ACGCCGCTGA TGGGCGAGTA CGCCGCGGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTT CTGCCCCAACG
1301 GTATGAAGCC CGAATACCGC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>:

```

a919.pep
1  MKKYLFRAL CGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SVQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGGDDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
201 HTADLSQFPI TARTTAIKGR FEGRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KAYMQQNPQR LAEVLGQNPS YIFFRELTGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap

```

                10      20      30      40      50      60
m919.pep      MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
                |||||
a919          MKKYLFRALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
                10      20      30      40      50      60

                70      80      90      100     110     120
m919.pep      YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
                |||||
a919          YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
                70      80      90      100     110     120

```

m919.pep	130	140	150	160	170	180
	YFTPWQVAGNGSLAGTVTGYYPVLKGD	DRRTAQA	RFFIYGIPDDFISVPLPAGLRSGKA			
a919	YFTPWQVAGNGSLAGTVTGYYPVLKGD	DRRTAQA	RFFIYGIPDDFISVPLPAGLRSGKA			
	130	140	150	160	170	180
m919.pep	190	200	210	220	230	240
	LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL					
a919	LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL					
	190	200	210	220	230	240
m919.pep	250	260	270	280	290	300
	DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSI	GRYMADKGYL				
a919	DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSI	GRYMADKGYL				
	250	260	270	280	290	300
m919.pep	310	320	330	340	350	360
	KLQQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA					
a919	KLQQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFREL	TGSSNDGPVGALGTPLMGEYAGA				
	310	320	330	340	350	360
m919.pep	370	380	390	400	410	420
	VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK					
a919	VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK					
	370	380	390	400	410	420
m919.pep	430	440				
	QKTTGYVWQLLPNGMKPEYRPX					
a919	QKTTGYVWQLLPNGMKPEYRPX					
	430	440				

121 and 121-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3075>:

m121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCG TACCCCGCA GGTACGCCG CCAATTGCTG
151 GATTTCGAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
401 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
451 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
501 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
551 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
601 xxxxxxCAGC TTCCTTACGA CAAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTGCGAC
701 AACGCCACCC TAAAAGCAGG GGGCGCGAAC TGTTTGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGCGCG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTTT ACCGCGCAAA CCGTTTGCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTTG
1001 CGTGGTTGGC GGC GTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA

```

1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101 A

This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>:

m121.pep

```

1  METQLYIGIM SGTSMGDGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRQLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL Axxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
151 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3077>:

g121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCT TACCCTGACC GGTTCGCCCG CAATGTCTG
151 GATTTCGAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACTGa cgcggatttT TACCGTCggc gacttcCGCA
401 GCCCGCACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCC GGCGCA CCGCCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
601 cactGGcgagc TGCCTTACGA CAAAacggt gcAAAGcgg cacAAGGCAA
651 catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
701 AACCccaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
751 gaaacctAcc ttgacggcg cgaaaaccga tacgacgtat tgcggacgt
801 ttcccgatcc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
851 CAGATGCCCG TCAAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001 cgtggttggC GGCGTGTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3078; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMGDGADA VLVRMDGGKW LGAEGHAFTP YPDRLLRRKLL
51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSONLAPC DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVWDAYS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDGADAVLIRMDGGKW LGAEGHAFTPYPGRLLRRQLLDLQDTGADEL					
	: : : : :					
g121	METQLYIGIMSGTSMGDGADAVLVRMDGGKW LGAEGHAFTPYPDRLLRRKLLDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEGYSIQLADLPLL					
	: : : : :					

```

g121      HRSRMLSQELSRLYAQTAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
           70      80      90      100      110      120
           130      140      150      160      170      180
m121.pep  AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
           | : : : : :
g121      AELTRIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDDRETRVVLNIGGIANISVLPPGA
           130      140      150      160      170      180
           190      200      210      220      230      240
m121.pep  XXXXXXXXXXXXXXXXXXXXXXXLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
           : : : |||||:||||| |||||:| |||||
g121      PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
           190      200      210      220      230      240
           250      260      270      280      290      300
m121.pep  GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICDGGIRNPV
           |||||:||||| ||||| ||||| ||||| ||||| ||||| |||||
g121      GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDVSHAAADARQMYICGGGIRNPV
           250      260      270      280      290      300
           310      320      330      340      350      360
m121.pep  LMADLAECFGTRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
           |||||:||||| ||||| ||||| ||||| ||||| ||||| |||||
g121      LMADLAECFGTRVSLHSTAE LNLDPOWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
           310      320      330      340      350      360

m121.pep  XAGYYYY
           |||||
g121      GAGYYYY

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3079>:

```

a121.seq
1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGGCGCGGC GGACAAAGCG CGCCGCTCGT CCGCGCTTT
451 CACGAAGCCC TGTTCGCGCA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCGCGACGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCGG TATTTGCGAC
701 AACCCACCCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCATTTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCC
1001 CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAEE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV

```


301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
 351 ATGASKPCIL GAGYYY*

m121/a121 ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
a121	METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL					
a121	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSVQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121.pep	XX					
a121	AERTQIFTVGDFRSRDLAAGGQGAPLVPAPFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
a121	PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICDGGIRNPV					
a121	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVFVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAE LNLDPQWVEAAAFAWMAACWVNRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3081>:

m121-1.seq

1	ATGGAACAC	AGCTTTACAT	CGGCATCATG	TCGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTAATGATAC	GGATGGACGG	CGGCAATGG	CTGGGCGCGG
101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA	GGTTACGCCG	CCAATTGCTG
151	GATTTGCAGG	ACACAGGCGC	AGACGAACTG	CACCGCAGCA	GGATTTTGTC
201	GCAAGAACTC	AGCCGCTAT	ATGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
251	GTCAAACCT	CGCACCGTCC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
301	ACCGTCCGAC	ACGCGCCGGA	ACACGGTTAC	AGCATAACGC	TGCGCGATTT
351	GCCGCTGCTG	GCGGAACGGA	CGCGGATTTT	TACCGTCGGC	GACTTCCGCA
401	GCCGCGACCT	TGCGGCCGCG	GGACAAGGCG	CGCCACTCGT	CCCCGCCTTT
451	CACGAAGCCC	TGTTCCGCGA	CAACAGGGAA	ACACGCGCGG	TACTGAACAT
501	CGGCGGGATT	GCCAACATCA	GCGTACTCCC	CCCCGACGCA	CCCGCCTTCG
551	GCTTCGACAC	AGGGCCGGGC	AATATGCTGA	TGGACGCGTG	GACGCAGGCA
601	CACTGGCAGC	TTCTTTACGA	CAAAAACGGT	GCAAAGGCGG	CACAAGGCAA
651	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	CGCCACCCG	TATTTGCGAC
701	AACCCACCC	TAAAAGCACG	GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC
751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA	TACGACGTAT	TGCGGACGCT
801	TTCCCGTTTT	ACCGCGCAAA	CCGTTTTCGA	CGCCGTCTCA	CACGCAGCGG

```

851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTTG
1001 CGTGGTTGGC GGCGTGTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3082; ORF 121-1>:

```

m121-1.pep
  1 METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRRLRRQLL
  51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
 101 TVRHAPHEGY SIQLADLPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPF
 151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWTQA
 201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
 251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICGGGIRNPV
 301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
 351 ATGASKPCIL XAGYYY*

```

m121-1/g121 ORFs 121-1 and 121.ng showed a 95.6% identity in 366 aa overlap

	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRRLRRQLLDLQDTGADEL					
	: : : : :					
g121	METQLYIGIMSGTSMGDADAVLVRMDGGKWLGAEGHAFTPYPDRLLRKLDDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121-1.pep	HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPHEGYSIQLADLPLL					
	: : : : :					
g121	HRSRMLSQELSRLYAQTAAELLCSONLAPCDITALGCHGQTVRHAPHEGYSIQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121-1.pep	AERTRIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDNRETRAVLNIGGIANISVLPPDA					
	: : : : :					
g121	AELTRIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDDRETRVVLNIGGIANISVLPPGA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121-1.pep	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	: : : : :					
g121	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAYSHAAADARQMYICGGGIRNPV					
	: : : : :					
g121	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDAYSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
	: : : : :					
g121	LMADLAECFGTRVSLHSTAE LNLDPQWVEAAAFWLAACWINRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
g121	GAGYYYYX					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3083>:

```

a121-1.seq
  1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG

```

```

51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGCGCCGGC GGACAAGGCG CGCCGCTCGT CCCCCTTTT
451 CACGAAGCCC TGTTCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGCGGGGATT GCCAACATCA GCGTACTCCC CCCCACGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCGG TATTTGCGAC
701 AACCACCCC TAAAGCACG GGGCGGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATT ACCGCGCAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3084; ORF 121-1.a>:

a121-1.pep

```

1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCQNLAAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPGP NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

m121-1.pep	10	20	30	40	50	60
	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
a121-1	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL					
	10	20	30	40	50	60
m121-1.pep	70	80	90	100	110	120
	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPHEHGYSIQLADLPLL					
a121-1	HRSRMLSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPESYSVQLADLPLL					
	70	80	90	100	110	120
m121-1.pep	130	140	150	160	170	180
	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA					
a121-1	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
m121-1.pep	190	200	210	220	230	240
	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTPGPNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121-1.pep	250	260	270	280	290	300
	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVASHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFVASHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300

	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLHSTADLNLDLPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAELNLDLPQWVEAAAFWMAACWVNRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
a121	GAGYYYYX					

128 and 128-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3085>:

```

m128.seq (partial)
  1 ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAACCCC GTTTTGATCA
 51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAAGTGGCGG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAACTC AACCAC
  1 TACGCCAGCG AAAAAGTGGC CGAAGCCAAA TACGCGTTCA GCGAAACCGA
 51 wGTCAAAAAA TAYTTCCCYG TCGGCAAwGT ATTAACCGGA CTGTTCGCCC
101 AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCCGTC
151 TGGCACAAGG ACGTGCCTTA TTKtGAATTG CAACAAAACG GCGAAmCCAT
201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
251 CGTGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
301 CAAYTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
351 CAGGGAAGCC CGCyTAGGCC ACGACGAAAT CCTCATCTC TTCCACGAAA
401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
451 TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
501 TATGGAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
551 ACGAAGAAAC CGGcgTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
601 GCCGCCAAAA ACTTCCAAsG CGGCATGTTC yTsGTCCGGC AAwTGGAGTT
651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCATC
751 CAGCCGCCCC AATACAACCG CTTGCCTTG AGCTTCGGC ACATCTTCGC
801 AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
851 GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
901 GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
951 nGCAGAAATC TTCAAAGCCT TCCGCGGCCG CGAACCGAGC ATAGACGCAC
1001 TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>:

```

m128.pep (partial)
  1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
 51 NTVEPLTGIT ERVGRWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH
//
  1 YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
 51 WHKDVRXYEL QONGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3087>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
51  aatccaaacc gaagACatca AACC CGCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACCGCTCG GCAGGATTGT
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
401 TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACC Gccaaa cTGCTCGGCT TAAAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTTCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCC
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTaca AAGGCCGCG CGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AacCGGCAC GGA CTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAacggcgtA GAATGGGACG CGGTGCAACT GCCCAGCCAG
1501 TTTATGGAAG ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TcgcCGCCAA AAATCTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCct
1851 cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgCGGAAT CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTGCTGCGC CAaagcggT TCGACAACGC gGcttgA

```

This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>:

```

g128.pep
1  MIDNALLHLG EEPFRNQIQT EDIKPAVQTA IAEARGQIAA VKAQHTTGWA
51  NTVERLTGIT ERVGRWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDAAAPLAGI PEDALAMFAA
201 AAGSEKGTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
351 EVKKYFPVVG VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQONGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS ESDECRLKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS

```

651 AAESFKAFRG REPSIDALLR QSGFDNAA*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng)
from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHGTGWTVERLTGIT					
	: :					
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHGTGWTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
g128.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA					
	:					
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130	140	150	160	170	180
g128.pep	TLSPAQKTKLDHDLRDFVLGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	:					
m128	TLSPAQKTKLNH					
	130					
	//					
				340	350	360
g128.pep	YAGEKLREAKYAFSETEVKKYFPVGKVLG					
	:					
m128	YASEKLREAKYAFSETXVKKYFPVGXVLNG					
		10	20	30		
	370	380	390	400	410	420
g128.pep	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWMNDYK					
	:					
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQONGEXIGGVYMDLYAREGKRGGAWMNDYK					
	40	50	60	70	80	90
	430	440	450	460	470	480
g128.pep	GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV					
	:					
m128	GRRRFSGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVDELGV					
	100	110	120	130	140	150
	490	500	510	520	530	540
g128.pep	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAANKFQGMF					
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAANKFQXGMF					
	160	170	180	190	200	210
	550	560	570	580	590	600
g128.pep	LVRQMEFALFDMMIYSEDECRLLKNWQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY					
	:					
m128	XVRQXEFALFDMMIYSEDEGRLLKNWQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY					
	220	230	240	250	260	270
	610	620	630	640	650	660
g128.pep	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAESFKAFRGREPS					
	: :					
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS					
	280	290	300	310	320	330

```

              670      679
g128.pep      IDALLRQSGFDNAAX
              |||||:|||||:
m128          IDALLRHSGFDNAVX
              340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3089>:

```

a128.seq
1   ATGACTGACA ACGCACTGCT CCATTGCGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCCTCG GCAGGATTGT
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAATC AACCACGATC
401 TGCGCGATTTC CGTCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAATTTCCTG CCAAATTCCT
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG GCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCGCGCCGCG CCAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCGCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGGCGCG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCGCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCAGTCAG
1501 TTTATGGAAT ATTTCTTTG GGAATACAAT GTCTTGCGCG AAATGTCCGC
1551 CCACGAAGAA ACCGCGGTTT CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCGGCCAA AAATCTCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CTTTGAAGA AAGCGACGAT GTCGCGGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

```

a128.pep
1   MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWAA
51  NTVEPLTGIT ERVGRWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDAAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
 601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS
 651 AAESFKAFRG REPSIDALLR HSGFDNAA*

m128/a128 ORFs 128 and 128.a showed a 66.0% identity in 677 aa overlap

	10	20	30	40	50	60
m128.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHGTWANTVEPLTGIT					
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHGTWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
m128.pep	ERVGRIGVGVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIGVGVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130					
m128.pep	TLSPAQKTKLNH-----					
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
				140	150	
m128.pep	-----YASEKLREAKYAFSETXVKKYFPVGX					
a128	ARRAKPYAEKDIAEVKAFARESGLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
	160	170	180	190	200	210
m128.pep	VLNGLFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
	220	230	240	250	260	270
m128.pep	NDYKGRRRFSDGTLQLPTAYLVCNFPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
	280	290	300	310	320	330
m128.pep	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540
	340	350	360	370	380	390
m128.pep	XGMFXVRQXEFALFDMMIYSEDDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600


```

m128.pep      400      410      420      430      440      450
               AGGYSAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRG
               |||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128          610      620      630      640      650      660
               AGGYSAAGYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG

m128.pep      460      470
               REPSIDALLRHSGFDNAVX
               ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128          670
               REPSIDALLRHSGFDNAAX

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3091>:

```

m128-1.seq
1   ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCG GAAGTGGCGG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAACTC AACCACGATC
401 TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 AACTACCTC GCCGTCTATC AATACGCCGA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCG GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCGCGCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGCGGCGG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCCGCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGCGG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG
1501 TTTATGGAAT ATTTCTGTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATCTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751 TCCAGCGGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3092; ORF 128-1>:

```

m128-1.pep.
1   MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWAA
51  NTVEPLTGIT ERVGRIGVVV SHLNSVADTP ELRAVYNELM PEITVFFTEI

```

```

101  GDIELYNRF  KTIKNSPEFD  TLSPAQKTKL  NHDLRDFVLS  GAELPPEQQA
151  ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201  AAQSESKTGY  KIGLQIPHYL  AVIQYADNRE  LREQIYRAYV  TRASELSDDG
251  KFDNTANIDR  TLANALQTAK  LLGFKNYAEL  SLATKMADTP  EQVLNFLHDL
301  ARRAKPYAEK  DLAEVKAFAR  ESLNLADLQP  WDLGYASEKL  REAKYAFSET
351  EVKKYFPVGK  VLNGLFAQIK  KLYGIGFTEK  TVPVWHKDVR  YFELQQNGET
401  IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFS  DGTLLQLPTAY  LVCNFAPPVG
451  GREARLSHDE  ILILFHETGH  GLHLLLTQVD  ELGVSGINGV  EWDAVELPSQ
501  FMENFVWEYN  VLAQMSAHEE  TGVPLPKELF  DKMLAAKNFQ  RGMFLVRQME
551  FALFDMMIYS  EDDEGRLLKNW  QQVLDSVRKK  VAVIQPPEYN  RFALSFGHIF
601  AGGYSAGYYS  YAWAEVLSAD  AYAAFEESDD  VAATGKRFWQ  EILAVGGSRS
651  AAESFKAFRG  REPSIDALLR  HSGFDNAV*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3093>:

g128-1.seq (partial)

```

1   ATGATTGACA  ACGCACTGCT  CCACTTGGGC  GAAGAACCCC  GTTTTAATCA
51  AATCAAAACC  GAAGACATCA  AATCCGCCGT  CCAAACCGCC  ATCGCCGAAG
101 CGCGCGGACA  AATCGCCGCC  GTCAAAGCGC  AAACGCACAC  CGGCTGGGCG
151 AACACCGTCG  AGCGTCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201 GGGCGTCGTG  TCCCATCTCA  ACTCCGTCGT  CGACACGCCC  GAACTGCGCG
251 CCGTCTATAA  CGAACTGATG  CCTGAAATCA  CCGTCTCTTT  CACCGAAATC
301 GGACAAGACA  TCGAACTGTA  CAACCGCTTC  AAAACCATCA  AAAATTCCCC
351 CGAATTTGCA  ACGCTTTCCC  CCGCACAAAA  AACCAAGCTC  GATCACGACC
401 TGCGCGATTT  CGTATTGAGC  GGCGCGGAAC  TGCCGCCCAG  ACGGCAGGCA
451 GAACTGGCAA  AACTGCAAAC  CGAAGGCGCG  CAACTTTCCG  CCAAATTCTC
501 CCAAAACGTC  CTAGACGCGA  CCGACGCGTT  CGGCATTTAC  TTTGACGATG
551 CCGCACCGCT  TGCCGGCATT  CCCGAAGACG  CGCTCGCCAT  GTTTGCCGCC
601 GCCGCGCAAA  GCGAAGGCAA  AACAGGTTAC  AAAATCGGCT  TGCAGATTCC
651 GCACTACCTT  GCCGTTATCC  AATACGCCGG  CAACCGCGAA  CTGCGCGAAC
701 AAATCTACCG  CGCCTACGTT  ACCCGTGCCA  GCGAACTTTC  AAACGACGGC
751 AAATTGACAA  ACACCGCCAA  CATCGACCGC  ACGCTCGAAA  ACGCATTGAA
801 AACCGCCAAA  CTGCTCGGCT  TTAAAAATTA  CGCCGAATTG  TCGCTGGCAA
851 CCAAAATGGC  GGACACGCCC  GAACAGGTTT  TAAACTTCCT  GCACGACTCT
901 GCCCGCCGCG  CCAAACCCTA  CGCCGAAAAA  GACCTCGCCG  AAGTCAAAGC
951 CTTCGCCCGC  GAACACCTCG  GTCTCGCCGA  CCCGCAGCCG  TGGGACTTGA
1001 GCTACGCCGG  CGAAAACTG  CGCGAAGCCA  AATACGCATT  CAGCGAAACC
1051 GAAGTCAAAA  AATACTTCCC  CGTCGGCAAA  GTTCTGGCAG  GCCTGTTCCG
1101 CCAAAATCAA  AAATCTACG  GCATCGGATT  CGCCGAAAAA  ACCGTTCCCG
1151 TCTGGCACAA  AGACGTGCGC  TATTTTGAAT  TGCAACAAAA  CGGCAAAACC
1201 ATCGGCGGCG  TTTATATGGA  TTTGTACGCA  CGCGAAGGCA  AACGCGGCGG
1251 CGCGTGATG  AACGACTACA  AAGGCCGCCG  CCGCTTTGCC  GACGGCACGC
1301 TGCAACTGCC  CACCGCTAC  CTCGTCTGCA  ACTTCGCCCC  GCCGTCGGC
1351 GGCAAAGAAG  CGCGTTTAAG  CCACGACGAA  ATCCTCAGCC  TCTTCCACGA
1401 AACCGGCCAC  GGAAGTCAAC  ACCTGCTTAC  CCAAGTGGAC  GAACTGGGCG
1451 TGTCGGGCAT  CAACGCGGTA  AAA

```

This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:

g128-1.pep (partial)

```

1   MIDNALLHLG  EEPFRNQIKT  EDIKPAVQTA  IAEARGQIAA  VKAQHTGWA
51  NTVERTLGIT  ERVGRIWGVV  SHLNSVVDTP  ELRAVYNELM  PEITVFFTEI
101 GDIELYNRF  KTIKNSPEFA  TLSPAQKTKL  DHDLRDFVLS  GAELPPERQA
151 ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201 AAQSEGKTGY  KIGLQIPHYL  AVIQYAGNRE  LREQIYRAYV  TRASELSNDG
251 KFDNTANIDR  TLENALKTAK  LLGFKNYAEL  SLATKMADTP  EQVLNFLHDL
301 ARRAKPYAEK  DLAEVKAFAR  EHLGLADPQP  WDLGYASEKL  REAKYAFSET
351 EVKKYFPVGK  VLAGLFAQIK  KLYGIGFAEK  TVPVWHKDVR  YFELQQNGKT
401 IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFA  DGTLLQLPTAY  LVCNFAPPVG
451 GKEARLSHDE  ILTLFHETGH  GLHLLLTQVD  ELGVSGINGV  K

```

m128-1/g128-1 ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa overlap

10 20 30 40 50 60

g128-1.pep	MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT
m128-1	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
	10 20 30 40 50 60
g128-1.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
m128-1	ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
	70 80 90 100 110 120
g128-1.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130 140 150 160 170 180
g128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
	190 200 210 220 230 240
g128-1.pep	TRASELSNDGKFDNTANIDRTLENALKTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAELGFKNYAELSLATKMADTPEQVLNFLHDL
	250 260 270 280 290 300
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDLGYASEKLREAKYAFSETEVKKYFPVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNADLPWDLGYASEKLREAKYAFSETEVKKYFPVGK
	310 320 330 340 350 360
g128-1.pep	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD
	430 440 450 460 470 480
g128-1.pep	ELGVSGINGVK
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3095>:

a128-1.seq

1	ATGACTGACA	ACGCACTGCT	CCATTGGGC	GAAGAACCCC	GTTTGTATCA
51	AATCAAAACC	GAAGACATCA	AACCCGCCCT	GCAAACCGCC	ATTGCCGAAG
101	CGCGCGAACA	AATCGCCGCC	ATCAAAGCCC	AAACGCACAC	CGGCTGGGCA
151	AACACTGTCTG	AACCCCTGAC	CGGCATCACC	GAACGCGTCG	GCAGGATTGT
201	GGGCGTGGTG	TCGCACCTCA	ACTCCGTCAC	CGACACGCCC	GAAGTGCAGC
251	CCGCCTACAA	TGAATTAATG	CCCGAAATTA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TCGAGCTGTA	CAACCGCTTC	AAAACCATCA	AAAACCTCCC
351	CGAGTTCGAC	ACCCTCTCCC	ACGCGCAAAA	AACCAAACTC	AACCACGATC
401	TGCGCGATTT	CGTCTCAGC	GGCGCGGAAC	TGCCGCCCGA	ACAGCAGGCA
451	GAATTGGCAA	AACTGCAAAC	CGAAGGCGCG	CAACTTTCCG	CCAAATTCTC

```

501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCCTG TGCCGGCATT CCGGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCCCGCG CCAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCCG GAAAGCCTCG GCCTCGCCGA TTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAAATCAAA AACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCTGCGACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGAAA ATTTGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGCGGTTT CCCTGCCGAA AGAACTCTTC GACAAATGC
1601 TCGCCGCCAA AACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTC
1801 GCAGGCGGCT ATTCGCGAGC CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CTTCAAAGC CTTCGCGGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTHTGWA
51 NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*

```

m128-1/a128-1 ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

```

          10      20      30      40      50      60
a128-1.pep MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAA IKAQHTHTGWANTVEPLTGIT
|||||
m128-1     MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAA IKAQHTHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90     100     110     120
a128-1.pep ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
|||||
m128-1     ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90     100     110     120

          130     140     150     160     170     180

```

a128-1.pep	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130 140 150 160 170 180
a128-1.pep	190 200 210 220 230 240
	FDDAAPLAGIPEDALAMFAAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
	190 200 210 220 230 240
a128-1.pep	250 260 270 280 290 300
	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL
	250 260 270 280 290 300
a128-1.pep	310 320 330 340 350 360
	ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLGYASEKLREAKYAFSETEVKKYFPVVGK
	310 320 330 340 350 360
a128-1.pep	370 380 390 400 410 420
	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
a128-1.pep	430 440 450 460 470 480
	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHLLTQVD
	430 440 450 460 470 480
a128-1.pep	490 500 510 520 530 540
	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540
a128-1.pep	550 560 570 580 590 600
	RGMFLVRQMEFALFDMMIYSEDDEGRKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF
m128-1	RGMFLVRQMEFALFDMMIYSEDDEGRKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIF
	550 560 570 580 590 600
a128-1.pep	610 620 630 640 650 660
	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
m128-1	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
	610 620 630 640 650 660
a128-1.pep	670 679
	REPSIDALLRHSGFDNAAX
m128-1	REPSIDALLRHSGFDNAVX
	670

```

1 ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
51 CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAATCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201 CTACAAATGG GGCAGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
251 TGATCAATT CGTTTACAAr AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCCCAGCACA TGGCGGCGGC AAGCCGsAAA ATCCCCGAcA GCCGcYTCAA
351 GGCCGCGGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401 ACGTCGGACT CTACATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TTTTACAGA ATGA

```

This corresponds to the amino acid sequence <SEQ ID 3098; ORF 206>:

m206.pep..

```

1 MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPR
101 ARDMAAASRK IPDSRXKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3099>:

g206.seq

```

1 atgttttccc ccgacaaaac ctttttcctc tgtctcggcg cactgctcct
51 cgcctcatgc ggcacgacct ccggcaaaca ccgccaaccg aaacccaaac
101 agacagtccg gcaaatccaa gccgtccgca tcagccacat cggccgcaca
151 caaggctcgc aggaatcat gctccacagc ctccggactca tcggcacgcc
201 ctacaaatgg ggcggcagca gcaccgcaac cggcttcgac tgcagcggca
251 tgattcaatt ggtttacaaa aacgccctca acgtcaagct gccgcgcacc
301 gcccgcgaca tggcggcggc aagccgcaaa atccccgaca gccgcctcaa
351 ggccggcgac atcgtattct tcaacaccgg cggcgcacac cgctactcac
401 acgtcggact ctacatcggc aacggcgaat tcatccatgc ccccggcagc
451 ggcaaaacca tcaaaaccga aaaactctcc acaccgtttt acgcaaaaaa
501 ctaccttga ggcatacgt tttttacaga atga

```

This corresponds to the amino acid sequence <SEQ ID 3100; ORF 206.ng>:

g206.pep

```

1 MFSPDKTLFL CLGALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIGRT
51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQLVYK NALNVKLPR
101 ARDMAAASRK IPDSRLKAGD IVFFNTGGAH RYSHVGLYIG NGEFIHAPGS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*

```

ORF 206 shows 96.0% identity over a 177 aa overlap with a predicted ORF (ORF 206.ng) from *N. gonorrhoeae*:

m206/g206

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
g206	MFSPDKTLFLCLGALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIGRTQGSQELMLHS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPR TARDMAAASRK IPDSRXKAGD					
g206	LGLIGTPYKW GGSSTATGFD CSGMIQLVYK NALNVKLPR TARDMAAASRK IPDSRLKAGD					
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS GKTIKTEKLSTPFYAKNYLGAHTFFTE					
	:					
g206	IVFFNTGGAH RYSHVGLYIG NGEFIHAPSS GKTIKTEKLSTPFYAKNYLGAHTFFTE					

130 140 150 160 170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3101>:

```
a206.seq
1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201 CTACAAATGG GCGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
251 TGATTCAATT CGTTTACAAA AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCGCGCGACA TGGCGGCGGC AAGCCGCAAA ATCCCCGACA GCCGCCTTAA
351 GGCCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGTACTCAC
401 ACGTCGGACT CTATATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAACCA TCAAACCGA AAAACTCTCC ACACCGTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

```
a206.pep
1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRP
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
a206	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRP TARDMAAASRK IPDSRLKAGD					
a206	LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRP TARDMAAASRK IPDSRLKAGD					
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS GKTIKTEKLS TPFYAKNYLGAHTFFTEX					
a206	LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS GKTIKTEKLS TPFYAKNYLGAHTFFTEX					
	130	140	150	160	170	

287

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3103>:

```
m287.seq
1  ATGTTTAAAC GCAGCGTAAT CGCAATGGCT TGTATTTTGT CCCTTTCAGC
51  CTGCGGGGGC GCGGGTGGCG GATCGCCCGA TGTCAAGTCG GCGGACACGC
101 TGTCAAACCC TGCCGCCCTT GTTGTCTCTG AAAAAGAGAC AGAGGCAAAG
151 GAAGATGCGC CACAGGCAGG TTCTCAAGGA CAGGGCGCGC CATCCGCACA
201 AGGCAGTCAA GATATGGCGG CGGTTTCGGA AGAAAATACA GGCAATGGCG
251 GTGCGGTAAC AGCGGATAAT CCCAAAAATG AAGACGAGGT GGCACAAAAT
301 GATATGCCGC AAAATGCCGC CCGTACAGAT AGTTCGACAC CGAATCACAC
351 CCCGGATCCG AATATGCTTG CCGGAAATAT GGAAAATCAA GCAACGGATG
401 CCGGGGAATC GTCTCAGCCG GCAAACCAAC CGGATATGGC AAATGCGGCG
451 GACGGAATGC AGGGGGACGA TCCGTCGGCA GGCGGGCAAA ATGCCGGCAA
501 TACGGCTGCC CAAGGTGCAA ATCAAGCCGG AAACAATCAA GCCGCCGGTT
551 CTTCAGATCC CATCCCCGCG TCAAACCTTG CACCTGCGAA TGGCGGTAGC
```

```

601 AATTTTGGAA GGGTTGATTT GGCTAATGGC GTTTTGATTG ACGGGCCGTC
651 GCAAAATATA ACGTTGACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
701 ATTTCTTGGA TGAAGAAGTA CAGCTAAAAT CAGAATTGTA AAAATTAAGT
751 GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
801 TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
851 TTATCTTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTTCT
901 GCACGGTCGA GCGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
951 TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001 ATTCGCGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051 GGGGCGGAAA AATTGCCCCG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
1101 ACCGGCAAAA GCGGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1151 TACTGCATTT CCATACGGAA AACGGCCGTC CGTACCCGAC CAGGGGCAGG
1201 TTTGCCGCAA AAGTCGATTT CGGCAGCAA TCTGTGGACG GCATTATCGA
1251 CAGCGCGCAT GATTTGCATA TGGGTACGCA AAAATTCAAA GCCGCCATCG
1301 ATGGAAACGG CTTTAAGGGG ACTTGACGG AAAATGGCAG CGGGGATGTT
1351 TCCGGAAGT TTTACGGCCC GGCCGGCGAG GAAGTGGCGG GAAAATACAG
1401 CTATGCCCCG ACAGATGCGG AAAAGGGCGG ATTCGGCGTG TTTGCCGGCA
1451 AAAAAGAGCA GGATTGA

```

This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

```

m287.pep
  1 MFKRSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVSEKETEAK
 51 EDAPQAGSQG QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
101 DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMANAA
151 DGMQGDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
201 NFRVLDLANG VLIDGPSQNI TLTHCKGDSG SGNNFLDEEV QLKSEFEKLS
251 DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS
301 ARSRRLPAE MPLIPVNQAD TLIVDGEAVS LTGHSNIFA PEGNYRYLTY
351 GAEKLPGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
401 FAAKVDFGSK SVDGIIDSGD DLHMGQTQFK AAIDGNGFKG TWTENGSGDV
451 SGKFYGPAGE EVAGKYSYRP TDAEKGFGV FAGKKEQD*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3105>:

```

g287.seq
  1 atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccctttcage
 51 ctgtgggggc ggcggtggcg gatcgcccga tgtcaagtcg gcggacacgc
101 cgtcaaaacc ggccgcccc gttgttgctg aaaatgccg ggaaggggtg
151 ctgccgaaag aaaagaaaga tgaggaggca gcgggcggtg cgccgcaagc
201 cgatacgcag gacgcaaccg ccggaagaag cagccaagat atggcgcgag
251 tttcggcaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
301 aaaaatgaag acgcgggggc gcaaaatgat atgccgcaa atgccgccga
351 atccgcaaat caaacaggga acaaccaacc cgccggttct tcagattccg
401 cccccgcgtc aaaccctgct cctgcgaatg gcggtagcga ttttggaag
451 acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaatataac
501 gttgaccac tgtaaaggcg attcttgtaa tgggtgataat ttattggatg
551 aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
601 attaaagcat ataaaaaaga cgagcaacgg gagaattttg tcggtttggt
651 tgctgacagg gtaaaaaagg atggaactaa caaatatata atcttctata
701 cggacaaacc acctactcgt tctgcacggt cgaggaggtc gttccggcc
751 gagattccgc tgattcccgt caatcaggcc gatacgtgta ttgtggatgg
801 ggaagcggtc agcctgacgg ggcattccg caatatcttc gcgccgaag
851 ggaattaccg gtatctgact tacggggcgg aaaaattgcc cggcggtatc
901 tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttgttgg
951 cacggccgtg tacaacggcg aagtgtgca tttccatatg gaaaacggcc
1001 gtccgtaccc gtccggaggc aggtttgccg caaaagtcga tttcggcagc
1051 aaatctgtgg acggcattat cgacagcggc gatgatttgc atatgggtac
1101 gcaaaaattc aaagccgcca tcgatgaaa cggctttaag gggacttggg
1151 cgaaaaatgg cgcggggat gtttcggaa ggttttacgg ccggccggcg
1201 gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggg
1251 cggattcggc gtgtttgccg gcaaaaaaga tcgggattga

```

This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>:

```

g287.pep
  1 MFKRSVIAMA CIFPLSACGG GGGGSPDVKS ADTPSKPAAP VVAENAGEGV

```


	10	20	30	40	49		
m287.pep	MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSE-----				KETEA		
g287	:	:	:	:	:		
	10	20	30	40	60		
m287.pep	50	60	70	80	90	100	109
	KEDAPQAGSQQGAPSAQGSQDMAAVSEENTGNNGGAVTADNPKNEDEVAQN	DMPQNAAGT					
g287	:	::		:			
	AGGAPQADTQD--ATAGEGSQDMAAVSAENTGNNGGAATTDNPKNEDAGAQN	DMPQNAA--					
	70	80	90	100	110		
m287.pep	110	120	130	140	150	160	169
	DSSTPNHTPDPNMLAGNMENQATDAGESSQPANQPDMANAADGMQGGDDPS	AGGQNAGNTA					
g287	-----						
m287.pep	170	180	190	200	210	220	229
	AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFRVLDLANGVLIDGPSQNI	TLTHCKGDS					
g287	:: :			:	:	:	
	-ESANQTGNNQPAGSSDSAPASNPAPANGGSDFGRTNVGN	SVVIDGPSQNI					
	120	130	140	150	160	170	
m287.pep	230	240	250	260	270	280	289
	CSGNNFLDEEVQLKSEFEKLSADAKISNYKKDGNKDFVGLVADSVQMKGIN	QYIIFYKP					
g287	: :		:	:	:		:
	CNGDNLLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGT	NKYIIFYTD					
	180	190	200	210	220	230	
m287.pep	290	300	310	320	330	340	349
	KPTSFARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHS	GNIFAPEGNYRYLT					
g287	:						
	KPPT-----RSARSRRSLPAEIPVNQADTLIVDGEAVSLTGHS	GNIFAPEGNYRYLT					
	240	250	260	270	280	290	
m287.pep	350	360	370	380	390	400	409
	YGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGR	FAAKVDFGS					
g287				:			
	YGAEKLPGGSYALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGR	FAAKVDFGS					
	300	310	320	330	340	350	
m287.pep	410	420	430	440	450	460	469
	KSVDGIIDSGDDLHMG	TQKFKAIDGN					
g287							
	KSVDGIIDSGDDLHMG	TQKFKAIDGN					
	360	370	380	390	400	410	
m287.pep	470	480	489				
	PTDAEKGFGVFAGKKEQDX						

g287 |||||:|
 PTDAEKGFGVFAGKKDRDX
 420 430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3107>:

a287.seq
 1 ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC
 51 CTGTGGGGGC GGCGGTGGCG GATCGCCCGA TGTTAAGTCG GCGGACACGC
 101 TGTCAAAACC TGCCGCCCTT GTTGTTACTG AAGATGTCGG GGAAGAGGTG
 151 CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC
 201 CGATACGCAG GACGCAACCG CCGGAAAAGG CCGTCAAGAT ATGGCGGCAG
 251 TTTCGGCAGA AAATACAGGC AATGGCGGTG CGGCAACAAC GGATAATCCC
 301 GAAAATAAAG ACGAGGGACC GCAAAATGAT ATGCCGCAAA ATGCCGCGCA
 351 TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA
 401 GAGATATGGG AAACCAAGCA CCGGATGCCG GGAATCGGC ACAACCGGCA
 451 AACCAACCGG ATATGGCAAA TGCGGCGGAC GGAATGCAGG GGGACGATCC
 501 GTCGGCAGGG GAAAATGCCG GCAATACGGC AGATCAAGCT GCAATCAAG
 551 CTGAAAACAA TCAAGTCGGC GGCTCTCAAA ATCCTGCCTC TTCAACCAAT
 601 CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA
 651 TGGCATCAAG CTTGACAGCG GTTCGGAAAA TGTAACGTTG ACACATTGTA
 701 AAGACAAAGT ATGCGATAGA GATTTCCTAG ATGAAGAAGC ACCACCAAAA
 751 TCAGAATTTG AAAAATTAAG TGATGAAGAA AAAATTAATA AATATAAAAA
 801 AGACGAGCAA CGAGAGAATT TTGTCGGTTT GGTGCTGAC AGGGTAGAAA
 851 AGAATGGAAC TAACAAATAT GTCATCATTT ATAAAGACAA GTCCGCTTCA
 901 TCCTCATCTG CGCGATTGAG GCGTTCGCA CCGTCGAGGC GGTGCGTTCC
 951 GGCCGAGATG CCGCTGATTC CCGTCAATCA GGCGGATACG CTGATTGTCT
 1001 ATGGGGAAGC GGTGAGCCTG ACGGGGCATT CCGGCAATAT CTTGCGGCC
 1051 GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCCGGCGG
 1101 ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG
 1151 CGGGCACGGC CGTGTACAAC GGCGAAGTGC TGCATTTCCA TATGGAAAAC
 1201 GGCCGTCCGT CCCCCTCCGG AGGCAGGTTT GCCGCAAAAG TCGATTTCGG
 1251 CAGCAAATCT GTGGACGGCA TTATCGACAG CGGCGATGAT TTGCATATGG
 1301 GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT
 1351 TGGACGGAAA ATGGCGGCGG GGATGTTTCC GGAAGGTTTT ACGGCCCGGC
 1401 CGGCGAAGAA GTGGCGGGA AATACAGCTA TCGCCGACA GATGCGGAAA
 1451 AGGGCGGATT CGGCGTGTTC GCCGGCAAAA AAGAGCAGGA TTGA

This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

a287.pep
 1 MFKRSVIAMA CIVALSAACGG GGGGSPDVKS ADTLSKPAAP VVTEDVGEEV
 51 LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATTDNP
 101 ENKDEGPOND MPQNAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAQPA
 151 NQPDMAAAD GMQDDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN
 201 PNATNGGSDF GRINVANGIK LDSGSENVTL THCKDKVCDR DFLDEEAPPK
 251 SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VIYKDKSAS
 301 SSSARFRRSA RSRRSLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP
 351 EGNRYRLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHFMEN
 401 GRPSPSGGRF AAKVDFGSKS VDGIIIDSGDD LHMGTQKFKA VIDGNGFKGT
 451 WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD*

m287/a287 ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

	10	20	30	40	49
m287.pep	MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----KETEA				
	:				
a287	MFKRSVIAMACIVALSAACGGGGGGSPDVKSADTLSKPAAPVVTEDEVGEEVLPKEKKDEEA				
	10	20	30	40	60
	50	60	70	80	100
m287.pep	KEDAPQAGSQGQAPSAGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMDMPQNAAGT				
	: ::: : : : : :				
a287	VSGAPQADTQ--DATAGKGGQDMAAVSAENTGNGGAATTDNPNENKDEGPONDMPQNAADT				
	70	80	90	100	110

m406.seq

8	ATGCAAGCAC	GGCTGCTGAT	ACCTATTCTT	TTTTCAGTTT	TTATTTTATC
51	CGCCTGCGGG	AACTGACAG	GTATTCCATC	GCATGGCGGA	GGTAAACGCT
101	TTGCGGTCTGA	ACAAGAACTT	GTGGCCGCTT	CTGCCAGAGC	TGCCGTTAA
151	GACATGGATT	TACAGGCATT	ACACGGACGA	AAAGTTGCAT	TGTACATTGC
201	CACATATGGGC	GACCAAGGTT	CAGGCAGATT	GACAGGGGGT	CGCTACTCCA
251	TTGATGCATC	GATTCGTGGC	GAATACATAA	ACAGCCCTGC	CGTCCGTACC
301	GATTACACCT	ATCCACGTTA	CGAAACCACC	GCTGAAACAA	CATCAGGCGG
351	TTTGACAGGT	TTAACCACTT	CTTTATCTAC	ACTTAATGCC	CCTGCACTCT
401	CTCGCACCCA	ATCAGACGGT	AGCGGAAGTA	AAAGCAGTCT	GGGCTTAAAT
451	ATTGGCGGGA	TGGGGGATTA	TCGAAATGAA	ACCTTGACGA	CTAACCCGCG
501	CGACACTGCC	TTTCTTTCCC	ACTTGGTACA	GACCGTATTT	TTCCTGCGCG
551	GCATAGACGT	TGTTTCTCCT	GCCAAATGCC	ATACAGATGT	GTTTATTAAC
601	ATCAGCAATC	TCGGAACGAT	ACGCAACAGA	ACCGAATGCT	ACCTATACAA
651	TGCCGGAATC	CTGAAAGCCC	AAACAAAACT	GGAATATTTC	GCAGTAGAC

```

701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACGGAAG GATTAATGGT CGATTTCTCC GATATCCGAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GTAGTGCGAC AACATAGACA
951 AGGACAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>:

m406.pep

```

1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGLTGT LTTSLSTLNA PALSRQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301 SHEGYGYSDE VVRQHRQGQP *

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3111>:

g406.seq

```

1  ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGACAG GTATTCCATC GCATGGCGGA GGCAAACGCT
101 TCGCGGTGCA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTGCGGCG GAATACATAA ACAGCCCTGC CGTCCGCACC
301 GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTGCA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT GTTTCTCCTT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAACT GGAATATTTT GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC AACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>:

g406.pep

```

1  MRARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGLTGT LTTSLSTLNA PALSRQSDG SGSRSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301 SHEGYGYSDE AVRQHRQGQP *

```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from *N. gonorrhoeae*:

g406/m406

	10	20	30	40	50	60
g406.pep	MRARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA.VKDMDLQALHGR					
	:					
m406	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA.VKDMDLQALHGR					

110

	10	20	30	40	50	60
	70	80	90	100	110	120
g406.pep	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
m406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
	70	80	90	100	110	120
	130	140	150	160	170	180
g406.pep	LTTSLSLTLNAPALSRTQSDGSGSRSSLGLNIGMGDYRNETLTNPRDTAFLSHLVQTVF					
m406	LTTSLSLTLNAPALSRTQSDGSGSKSSLGLNIGMGDYRNETLTNPRDTAFLSHLVQTVF					
	130	140	150	160	170	180
	190	200	210	220	230	240
g406.pep	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
m406	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	190	200	210	220	230	240
	250	260	270	280	290	300
g406.pep	IKPKTNAFEAAAYKENYALWMGPYKVS KGIKPTEGLMVD FSDIQPYGNHTGNSAPSVEADN					
m406	IKPKTNAFEAAAYKENYALWMGPYKVS KGIKPTEGLMVD FSDIRPYGNHTGNSAPSVEADN					
	250	260	270	280	290	300
	310	320				
g406.pep	SHEGYGYSDEAVRQHRQGQPX					
m406	SHEGYGYSDEVVRQHRQGQPX					
	310	320				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3113>:

```

a406.seq
1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG ACACTGCAGAG GTATTCCATC GCATGGCGGA GGTAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCAACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATTGCGGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTCCTCCT GCCAATGCCG ATACGGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACCT GGAATATTTC GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGACCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACAGAAG GATTAATGGT CGATTCTCTC GATATCCAAC
851 CATACGGCAA TCATATGGGT AACTCTGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGATA CAGCGATGAA GCAGTGCGAC GACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

```

a406.pep
1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARA AVK
51  DMDLQALHGR KVALYIATMG DQSGSGSLTG RYSIDALIRG EYINSPAVRT
101 DYTPRYETT AETTSGLTG LTTSLSLTLN PALSRQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVS PANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA

```

```

251 AYKENYALWM GPYKVSKEGIK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN
301 SHEGYGYSDE AVRRHRQGQP *

m406/a406 ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap

      10      20      30      40      50      60
m406.pep MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAANKDMDLQALHGR
|||||
a406      MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAANKDMDLQALHGR
      10      20      30      40      50      60

      70      80      90     100     110     120
m406.pep KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGTG
|||||
a406      KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGTG
      70      80      90     100     110     120

      130     140     150     160     170     180
m406.pep LTTSLSLTNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTNPRDTAFLSHLVQTVF
|||||
a406      LTTSLSLTNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTNPRDTAFLSHLVQTVF
      130     140     150     160     170     180

      190     200     210     220     230     240
m406.pep FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
|||||
a406      FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL
      190     200     210     220     230     240

      250     260     270     280     290     300
m406.pep IKPKTNAFEAAAYKENYALWMGPYKVSKEGIKPTGLMVDFS DIRPYGNHTGNSAPSVEADN
|||||
a406      IKPKTNAFEAAAYKENYALWMGPYKVSKEGIKPTGLMVDFS DIQPYGNHMGNSAPSVEADN
      250     260     270     280     290     300

      310     320
m406.pep SHEGYGYSDEVVRQHRQGQPX
|||||
a406      SHEGYGYSDEAVRRHRQGQPX
      310     320

```

EXAMPLE 2

Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene 919 was cloned in pET vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm

that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 3

Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene 279 was cloned in pGex vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 4

Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene 576 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera

were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B).. These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 5

Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene 519 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 6

Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene *121* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 7

Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene *128* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J.*

Immunol 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 8

Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene 206 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worth noting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in *E. coli* without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from *E. coli* expressing this native form of the 206 protein showed a reactive band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 206 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 9

Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene 287 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 10

Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene 406 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al.

1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 11

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

Table 2

225 gene variability: List of used *Neisseria* strains

Identification Strains number	Source / reference
Group B	
zo01_225 NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo02_225 BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zo03_225 NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo04_225 297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zo05_225 1000	R. Moxon / Seiler <i>et al.</i> , 1996
zo06_225 BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zo07_225 BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zo08_225 528	R. Moxon / Seiler <i>et al.</i> , 1996
zo09_225 NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zo10_225 BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zo11_225 NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zo12_225 NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zo13_225 NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zo14_225 NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zo15_225 SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zo16_225 NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zo17_225 NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zo18_225 BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zo19_225 BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zo20_225 44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zo21_225 MC58	R. Moxon
zo96_225 2996	Our collection
Group A	
zo22_225 205900	R. Moxon
zo23_225 F6124	R. Moxon
z2491 Z2491	R. Moxon / Maiden <i>et al.</i> , 1998
Group C	
zo24_225 90/18311	R. Moxon
zo25_225 93/4286	R. Moxon

Others

zo26_225 A22 (group W) R. Moxon / Maiden *et al.*, 1998
 zo27_225 E26 (group X) R. Moxon / Maiden *et al.*, 1998
 zo28_225 860800 (group Y) R. Moxon / Maiden *et al.*, 1998
 zo29_225 E32 (group Z) R. Moxon / Maiden *et al.*, 1998

Gonococcus

zo32_225 Ng F62 R. Moxon / Maiden *et al.*, 1998
 zo33_225 Ng SN4 R. Moxon

 fa1090 FA1090 R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

>FA1090 <SEQ ID 3115>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
 GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
 RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN*

Z2491 <SEQ ID 3116>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA
 DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
 MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
 IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN*

Z001_225 <SEQ ID 3117>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
 VKKNDSRFLN*

Z002_225 <SEQ ID 3118>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
 VKKNDSRFLN*

Z003_225 <SEQ ID 3119>
 MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
 VKKNDSRFLN*

Z004_225 <SEQ ID 3120>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z005_225 <SEQ ID 3121>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z006_225 <SEQ ID 3122>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z007_225 <SEQ ID 3123>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z008_225 <SEQ ID 3124>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z009_225 <SEQ ID 3125>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z010_225 <SEQ ID 3126>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z011_225 <SEQ ID 3127>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRFI
HAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

Z012_225 <SEQ ID 3128>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

Z013_225 <SEQ ID 3129>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z014_225 <SEQ ID 3130>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z015_225 <SEQ ID 3131>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKFRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z016_225 <SEQ ID 3132>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z017_225 <SEQ ID 3133>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z018_225 <SEQ ID 3134>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z019_225 <SEQ ID 3135>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z020_225 <SEQ ID 3136>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPINRAPARRAGNADELIGSAMGLNEQPVLVPVNRVPARRAGNA
DELIGNAMGLNEQPVLVPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKFRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRFI
HAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z021_225 <SEQ ID 3137>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z022_225 <SEQ ID 3138>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR

SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z023_225 <SEQ ID 3139>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z024_225 <SEQ ID 3140>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z025_225 <SEQ ID 3141>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z026_225 <SEQ ID 3142>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z027_225 <SEQ ID 3143>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z028_225 <SEQ ID 3144>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z029_225 <SEQ ID 3145>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z032_225 <SEQ ID 3146>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIKFRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z033_225 <SEQ ID 3147>

MDSFFKPAVWAVLWLMFAVRSALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIKFRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z096_225 <SEQ ID 3148>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA
 DELIGNAMGLLGIAYRYGGTSISTGFDSCGFMQHIKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHPRTGKNIEITSLSHKYWSGKYAFARR
 VKKNDPSRFLN*

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 12

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

Table 3

235 gene variability: List of used *Neisseria* strains

Identification Strains number	Reference
Group B	
gnmzq01 NG6/88	Seiler <i>et al.</i> , 1996
gnmzq02 BZ198	Seiler <i>et al.</i> , 1996
gnmzq03 NG3/88	Seiler <i>et al.</i> , 1996
gnmzq04 1000	Seiler <i>et al.</i> , 1996
gnmzq05 1000	Seiler <i>et al.</i> , 1996
gnmzq07 BZ169	Seiler <i>et al.</i> , 1996
gnmzq08 528	Seiler <i>et al.</i> , 1996
gnmzq09 NGP165	Seiler <i>et al.</i> , 1996
gnmzq10 BZ133	Seiler <i>et al.</i> , 1996
gnmzq11 NGE31	Seiler <i>et al.</i> , 1996
gnmzq13 NGE28	Seiler <i>et al.</i> , 1996
gnmzq14 NGH38	Seiler <i>et al.</i> , 1996
gnmzq15 SWZ107	Seiler <i>et al.</i> , 1996
gnmzq16 NGH15	Seiler <i>et al.</i> , 1996
gnmzq17 NGH36	Seiler <i>et al.</i> , 1996
gnmzq18 BZ232	Seiler <i>et al.</i> , 1996
gnmzq19 BZ83	Seiler <i>et al.</i> , 1996
gnmzq21 MC58	Virji <i>et al.</i> , 1992
Group A	
gnmzq22 205900	Our collection

gnmzq23	F6124	Our collection
z2491	Z2491	Maiden <i>et al.</i> , 1998

Group C

gnmzq24	90/18311	Our collection
gnmzq25	93/4286	Our collection

Others

gnmzq26	A22	(group W) Maiden <i>et al.</i> , 1998
gnmzq27	E26	(group X) Maiden <i>et al.</i> , 1998
gnmzq28	860800	(group Y) Maiden <i>et al.</i> , 1998
gnmzq29	E32	(group Z) Maiden <i>et al.</i> , 1998
gnmzq31		<i>N. lactamica</i> Our collection

Gonococcus

gnmzq32	Ng F62	Maiden <i>et al.</i> , 1998
gnmzq33	Ng SN4	Our collection

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
 Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
 Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3149>
 MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
 AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ01 <SEQ ID 3150>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ02 <SEQ ID 3151>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ03 <SEQ ID 3152>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST

AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ04 <SEQ ID 3153>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ05 <SEQ ID 3154>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ07 <SEQ ID 3155>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ08 <SEQ ID 3156>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ09 <SEQ ID 3157>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AEPSEAGYVFPAAVVEETFKQNGLTNAADIHAVQPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ10 <SEQ ID 3158>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ11 <SEQ ID 3159>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ13 <SEQ ID 3160>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ14 <SEQ ID 3161>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ15 <SEQ ID 3162>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ16 <SEQ ID 3163>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ17 <SEQ ID 3164>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ18 <SEQ ID 3165>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ19 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ21 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ22 <SEQ ID 3167>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ23 <SEQ ID 3168>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ24 <SEQ ID 3169>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ25 <SEQ ID 3170>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ26 <SEQ ID 3171>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ27 <SEQ ID 3172>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

ins. Dark
on of amino
nservation
for both

ervation of



Group B

287_2	BZ198	Seiler <i>et al.</i> , 1996
287_9	NGP165	Seiler <i>et al.</i> , 1996
287_14	NGH38	Seiler <i>et al.</i> , 1996
287_21	MC58	Virji <i>et al.</i> , 1992

Group A

z2491	Z2491	Maiden <i>et al.</i> , 1998
-------	-------	-----------------------------

Gonococcus

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

References:

- Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

287_14 <SEQ ID 3179>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGGQDMAAVSEENTGNNGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS
NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ
TAGSQNPASSTNPSATNSGGDFGRITNVGNSVVIDGPSQNTLTHCKGDS CSGNNFLDEEV
QLKSEFEKLSADADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTS FAR
FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
DSGDGLHMGTTQKFKAAIDGNGFKGTWTENG GGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
GFGVFAGKKEQD*

287_2 <SEQ ID 3180>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGGQDMAAVSEENTGNNGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS
NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ
TAGSQNPASSTNPSATNSGGDFGRITNVGNSVVIDGPSQNTLTHCKGDS CSGNNFLDEEV
QLKSEFEKLSADADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTS FAR
FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
DSGDGLHMGTTQKFKAAIDGNGFKGTWTENG GGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
GFGVFAGKKEQD*

287_21. <SEQ ID 3181>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNNGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMAANTADGMQGGDDPSAGGQAGNTAAQGANQAGNNQ
AAGSSDPI PASNPAPANGGSNFGRVDLANGVLDGPSQNTLTHCKGDS CSGNNFLDEEV
QLKSEFEKLSADADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTS FARFRRS

ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVVDGIIDSGD
DLHMGTOQKFKAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV
FAGKKEQD*

287_9 <SEQ ID 3182>

MFKRSVIAMACIVALSACGGGGGGSPDVKSADTLSPKPAAPVVTEDEVGEEVLPKEKKDEEA
VSGAPQADTQDATAGKGGQDMAAVSAENTGNGGAATTDPENKDEGPQNMPQNAADTDS
STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADMGGDDPSAGENAGNTADQA
ANQAENNQVGGSQNPASSTNPATNGGSDFGRIINVANGIKLDSGSENVTLTHCKDKVCDR
DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVEKNGTNKYVIIYKDKSAS
SSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYG
AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKS
VDGIIDSGDDLHMGTOQKFKAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPT
DAEKGFGVFAGKKEQD*

FA1090 <SEQ ID 3183>

MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTSPKPAAPVVAENAGEGVLPKEKKDEEA
AGGAPQADTQDATAGEGSQDMAAVSAENTGNGGAATTDPKNEDAGAQNMPQNAEASAN
QTGNNQFAGSSDSAPASNAPANGGSDFGRTNVGNSVVIDGPSQNIITLTHCKGDSNGDN
LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVEKNGTNKYIIFYTDKPPT
SARSRRSLPAEIPVQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLPGGS
YALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKSVVDGIIDSG
DDLHMGTOQKFKAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPTDAEKGFGV
VFAGKKDRD*

Z2491 <SEQ ID 3184>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDVAQNMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMANAADMGGDDPSAGGQNAAGNTAAQGANQAGNNQ
AAGSSDPIPASNPAPANGGSGNFRVLDLANGVLIDGPSQNIITLTHCKGDSNGNNFLDEEV
QLKSEFEKLSADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPTSFARFRRS
ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVVDGIIDSGD
DLHMGTOQKFKAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV
FAGKKEQD*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 14

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

Table 5

519 gene variability: List of used *Neisseria* strains

Identification Strains	Source / reference
------------------------	--------------------

number**Group B**

zv01_519	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv02_519	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zv03_519ass	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv04_519	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zv05_519	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zv06_519ass	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zv07_519	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zv11_519	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zv12_519	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zv18_519	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zv19_519	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zv20_519ass	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection

Group A

zv22_519ass	205900	R. Moxon
z2491_519	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

Others

zv26_519	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zv27_519	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zv28_519	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zv29_519ass	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998

Gonococcus

zv32_519	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
fa1090_519	FA1090	R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090_519 <SEQ ID 3185>
MEFFIIILLAAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

Z2491_519 <SEQ ID 3186>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV01_519 <SEQ ID 3187>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV02_519 <SEQ ID 3188>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV03_519 <SEQ ID 3189>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV04_519 <SEQ ID 3190>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV05_519 <SEQ ID 3191>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV06_519ASS <SEQ ID 3192>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVFSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERK
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV07_519 <SEQ ID 3193>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV11_519 <SEQ ID 3194>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV12_519 <SEQ ID 3195>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV18_519 <SEQ ID 3196>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV19_519 <SEQ ID 3197>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV20_519ASS <SEQ ID 3198>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM
ISAGMKIIDSSKTAK*

ZV21_519ASS <SEQ ID 3199>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV22_519ASS <SEQ ID 3200>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV26_519 <SEQ ID 3201>

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTD PKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV27_519 <SEQ ID 3202>

```
MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV28_519 <SEQ ID 3203>

```
MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV29_519ASS <SEQ ID 3204>

```
MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSIVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREPEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSNTAK*
```

ZV32_519 <SEQ ID 3205>

```
MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV96_519 <SEQ ID 3206>

```
MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 15

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

Table 6

919 gene variability: List of used *Neisseria* strains

Identification Strains	Source / reference
------------------------	--------------------

number**Group B**

zm01	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm02	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zm03	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm04	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zm05	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zm06	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zm07	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zm08n	528	R. Moxon / Seiler <i>et al.</i> , 1996
zm09	NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zm10	BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zm11asbc	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zm12	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zm13	NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zm14	NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zm15	SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zm16	NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zm17	NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zm18	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zm19	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zm20	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection

Group A

zm22	205900	R. Moxon
zm23asbc	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

Group C

zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon

Others

zm26	A22 (group W)	R. Moxon / Maiden <i>et al.</i> , 1998
zm27bc	E26 (group X)	R. Moxon / Maiden <i>et al.</i> , 1998
zm28	860800 (group Y)	R. Moxon / Maiden <i>et al.</i> , 1998
zm29asbc	E32 (group Z)	R. Moxon / Maiden <i>et al.</i> , 1998
zm31asbc	<i>N. lactamica</i>	R. Moxon

Gonococcus

zm32asbc	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
zm33asbc	Ng SN4	R. Moxon

fa1090	FA1090	R. Moxon
--------	--------	----------

References:

- Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>

MKKHLIRSALYGIAAAILAACQSRSIQTFFQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRTERARFPIYGIPTDDFISVPLPAGLRGGKN
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

Z2491 <SEQ ID 3208>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM01 <SEQ ID 3209>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM02 <SEQ ID 3210>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM03 <SEQ ID 3211>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM04 <SEQ ID 3212>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM05 <SEQ ID 3213>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM06 <SEQ ID 3214>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM07 <SEQ ID 3215>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM08N <SEQ ID 3216>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM09 <SEQ ID 3217>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM10 <SEQ ID 3218>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTGTGYEPEVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
 LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL

KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM11ASBC <SEQ ID 3219>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM12 <SEQ ID 3220>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM13 <SEQ ID 3221>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM14 <SEQ ID 3222>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSRNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM15 <SEQ ID 3223>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D L A G T T V G G G G A V
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM16 <SEQ ID 3224>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPGRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM17 <SEQ ID 3225>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM18 <SEQ ID 3226>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM19 <SEQ ID 3227>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM20 <SEQ ID 3228>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM21 <SEQ ID 3229>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM22 <SEQ ID 3230>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM23ASBC <SEQ ID 3231>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL

KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM24 <SEQ ID 3232>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM25 <SEQ ID 3233>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM26 <SEQ ID 3234>

MKKYLFRAALYGISAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKAYMQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM27BC <SEQ ID 3235>

MKKYLFRAALYGISAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM28 <SEQ ID 3236>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM29ASBC <SEQ ID 3237>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRI RQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATTHPITRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM31ASBC <SEQ ID 3238>

MKKHLFRAALYGIAAAILAACQSKSIQTFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYIGIPDDFISVPLPAGLRSGKA
 LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYVFFRELAGSGNDGPVGALGTPLMGEYAGA
 VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM32ASBC <SEQ ID 3239>

MKKHLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYIGIPDDFISVPLPAGLRGGKA
 LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGGDPVGALGTPLMGGYAGA
 IDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM33ASBC <SEQ ID 3240>

MKKHLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYIGIPDDFISVPLPAGLRGGKN
 LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
 IDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM96 <SEQ ID 3241>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYIGIPDDFISVPLPAGLRSGKA
 LVRIQRTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
 VDRHYITLGAPLFFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 16

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 7: Oligonucleotides used for PCR to amplify complete or partial ORFs

ORF	primer	Sequence	Restriction sites
001	Forward	CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGTCTTGCCAC	XhoI
003	Forward	CGCGGATCCCATATG-GTCGTATTCGTGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAATCATGAACACGCGC	XhoI
005	Forward	CGCGGATCCCATATG-GACAATATTGACATGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATCACATCCGCCG	XhoI
006	Forward	CGCGGATCCCATATG-CTGCTGGTGCTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGTTCCGGCTTTGATGT	XhoI
007	Forward	CGCGGATCCCATATG-GCCGACAACAGCATCAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAGGCGTTCATGATATAAG	XhoI
008	Forward	CGCGGATCCCATATG-AACAACAGACATTTTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCTGTCCGGTAAAAGAC	XhoI
009	Forward	CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGGCTTTTGCCACGTTTT	XhoI
011	Forward	CGCGGATCCCATATG-AAGACACACCGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGGTCAGTACGGT	XhoI
012	Forward	CGCGGATCCCATATG-CTCGCCCGTTGCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCGGGGAAGAGGCAC	XhoI
013	Forward	CGCGGATCCCATATG-CCTTTGACCATGCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGATTCGGCAAAAAATCT	XhoI
018	Forward	CGCGGATCCCATATG-CAGCAGAGGCAGTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGAGGCGAACGCC	XhoI
019	Forward	AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	Eco RI
	Reverse	AAACTGCAG-TCAGCGGGCGGGGACAATGCCCAT	Pst I
023	Forward	AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
	Reverse	AAACTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025	Forward	AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
	Reverse	AAACTGCAG-TCAGAACGCGATATAGCTGTTCCG	Pst I
031	Forward	CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATGTAAGACGGGGACAAC	XhoI
032	Forward	CGCGGATCCCATATG-CGGCGAAACGTGC	BamHI-

	Reverse	CCCGCTCGAG-CTGGTTTTTTGATATTTGTG	NdeI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	XhoI
	Reverse	CCCGCTCGAG-ATTTGCCGCATCCCGAT	BamHI-
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	NdeI
	Reverse	CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-
	Reverse	CCCGCTCGAG-CCGGA CTGCGTATCGG	NdeI
038	Forward	CGCGGATCCCATATG-ACCGATTTCCGCCA	XhoI
	Reverse	CCCGCTCGAG-TTCTACGCCGTACTGCC	BamHI-
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	NdeI
	Reverse	CCCGCTCGAG-TAGGATGACGAGGTAGG	XhoI
041	Forward	CGCGGATCCCATATG-TTCGTGCGCGAACCGC	BamHI-
	Reverse	CCCGCTCGAG-GCCCAAAA ACTCTTTCAA	NdeI
042	Forward	CGCGGATCCCATATG-ACGATGATTTGCTTGC	XhoI
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	BamHI-
043	Forward	AAAAAAGGTACC-ATGGTTGTTTCAAATCAAATATC	NdeI
	Reverse	AAACTGCAG-TTATTGCGCTTACCTTCCGCCGC	XhoI
043a	Forward	AAAAAAGGTACC-GCAAAAGTGCATGGCGGCTTGGACGGTGC	Kpn I
	Reverse	AAAAAACTGCAG-TTAATCCTGCAACACGAATTCGCCCCGTCCG	Pst I
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	Kpn I
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	Pst I
046	Forward	AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	BamHI-
	Reverse	AAACTGCAG-TCACTCGGCGACCCACACCGTGAA	NdeI
047	Forward	CGCGGATCCCATATG-GTCATCATA CAGGCG	XhoI
	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	Eco RI
048	Forward	AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	Pst I
	Reverse	AAACTGCAG-TCAAGATTCGACGGGGATGATGCC	Eco RI
049	Forward	AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Pst I
	Reverse	AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	BamHI-
050	Forward	CGCGGATCCCATATG-GGCGCGGGCTGG	NdeI
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	XhoI
052	Forward	AAAAAAGAATTC-ATGGCTTTGGTGGCGGAGGAAAC	Eco RI
	Reverse	AAAAAAGTCGAC-TCAGGCGGCGTTTTTACCTTCCT	Sal I
052a	Forward	AAAAAAGAATTC-GTGGCGGAGGAAACGGAAATATCCGC	Eco RI

	Reverse	AAAAAACTGCAG-TTAGCTGTTTTTGGAAACGCCGTCCAACCC	Pst I
073	Forward	CGCGGATCCCATATG-TGTATGCCATATAAGAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACCGGATTGTCCGAC	XhoI
075	Forward	CGCGGATCCCATATG-CCGTCTTACTTCATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCACCAATGCCGATTATTT	XhoI
077a	Forward	AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	Pst I
080	Forward	AAAGAATTC-GCGTCCGGGCTGGTTTGGTTTACAATTC	Eco RI
	Reverse	AAACTGCAG-CTATTCTTCGGATTCTTTTCGGG	Pst I
081	Forward	AAAGAATTC-ATGAAACCACTGGACCTAAATTTTCATCTG	Eco RI
	Reverse	AAACTGCAG-TCACTTATCCTCCAATGCCTC	Pst I
082	Forward	AAAGAATTC-ATGTGGTTGTTGAAGTTGCCTGC	Eco RI
	Reverse	AAACTGCAG-TTACGCGGATTCGGCAGTTGG	Pst I
084	Forward	AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
	Reverse	AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward	CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	XhoI
086	Forward	AAAAAAGGTACC-TATTTGGCATCAAAAGAAGGCGG	Kpn I
	Reverse	AAACTGCAG-TTACTCCACCCGATAACCGCG	Pst I
087	Forward	AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Eco RI
	Reverse	AAACTGCAG-TTACGCCGCACACGCAATCGC	Pst I
087a	Forward	AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCTGCAAGATGCCAGCTTGCC	Pst I
088	Forward	AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	Pst I
089	Forward	CGCGGATCCCATATG-CCGCCAAAATCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCGCATACCAAAGCCA	XhoI
090	Forward	CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAAAACGGCGGTACG	XhoI
091	Forward	AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	Eco RI
	Reverse	AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Pst I
092	Forward	AAAGAATTC-ATGTTTTTTATTTCAATCCG	Eco RI
	Reverse	AAACTGCAG-TCAAATCTGTTTCGACAATGC	Pst I
093	Forward	AAAGAATTC-ATGCAGAATTTTGGCAAAGTGGC	Eco RI
	Reverse	AAACTGCAG-CTATGGCTCGTCATACCGGGC	Pst I
094	Forward	AAAGAATTC-ATGCCGTACGGAAGCGCATCAACTC	Eco RI
	Reverse	AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Pst I
095	Forward	AAAGAATTC-ATGTCCTTTTCATTTGAACATGGACGG	Eco RI
	Reverse	AAACTGCAG-TCAACGCCGAGGCACTAACGCCC	Pst I
096	Forward	AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Eco RI

	Reverse	AAACTGCAG-TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097	Forward	AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
	Reverse	AAACTGCAG-TCAGCCCAAATACCAGAATTTTCAG	Pst I
098	Forward	AAAGAATTC-GATGAACGCAGCCCAGCATGGATACG	Eco RI
	Reverse	AAACTGCAG-TTACGACATTCTGATTTGGCA	Pst I
102	Forward	AAAAAAGAATTC-GGCCTGATGATTTTGGGAAGTCAACAC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward	CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGTTCTGCCAGTTTCAG	XhoI
107	Forward	AAAAAAGAATTC- CTGATGATTTTGGGAAGTCAACACCCATTATCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
107b	Forward	AAAAAAGAATTC- GATACCCAAGCCCCCGCCGGCACAACACTACTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC	Pst I
108	Forward	AAAGAATTC-ATGTTGCCGGGCTTCAACCG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
108a	Forward	AAAAAAGAATTC-GGTAACACATTCGGCAGCTTAGACGGTGG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
109	Forward	AAAGAATTC-ATGTATTATCGCCGGGTATGGG	Eco RI
	Reverse	AAACTGCAG-CTAGCCCAAAGATTTGAAGTGTC	Pst I
111	Forward	CGCGGATCCCATATG-TGTTCCGAACAAACCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGAGCAGTTTTTCAAA	XhoI
114	Forward	CGCGGATCCCATATG-GCTTCCATCACTTCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATCCGCGAAATCGTC	XhoI
117	Forward	AAAAAAGGTACC-ATGGTCGAAGAACTGGAAGTCTG	Kpn I
	Reverse	AAACTGCAG-TTAAAGCCGGGTAAACGCTCAATAC	Pst I
118	Forward	AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Sal I
	Reverse	AAAGCATGC-CTATTTTTTGTGTAATAATCAAATC	Sph I
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC	XhoI
122	Forward	CGCGGATCCCATATG-GTCATGATTAAATCCGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATCTTGGTAGATTGGATTT	XhoI
125	Forward	AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse	AAACTGCAG-TCACGCCGTTTCAAGACG	Pst I
125a	Forward	AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse	AAAAAACTGCAG- TTATTTTGCCACGTCGGTTTCTCCGGTGAACAACGC	Pst I
126	Forward	CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

	Reverse	CCCGCTCGAG-ATATTCCGCCGAATGCC	NdeI
127	Forward	AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG	XhoI
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Eco RI
127a	Forward	AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCGGG	Pst I
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Eco RI
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT	Pst I
	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA	BamHI-
130	Forward	CGCGGATCCCATATG-AAACAACCTCCGCGA	NdeI
	Reverse	CCCGCTCGAG-GAATTTTGACCCGGATTG	XhoI
132	Forward	AAAGAATTC-ATGGAACCTTCAAACCTTAATTTG	Eco RI
	Reverse	AAAAAACTGCAG-TCACCATGTCGGCATTGAAAAAC	Pst I
134	Forward	CGCGGATCCCATATG-TCCCAAGAAATCCTC	BamHI-
	Reverse	CCCGCTCGAG-CAGTTTGACCGAATGTTC	NdeI
135	Forward	CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	XhoI
	Reverse	CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	BamHI-
137	Forward	AAAAAAGGTACC-ATGATTACCCATCCCAATTCGATCC	NdeI
	Reverse	AAAAAACTGCAG-TCAGTGCTGTTTTTTCATGCCGAA	XhoI
137a	Forward	AAAAAAGAATTC-GGCCGAAACACGGCATCGGCTTCCT	Kpn I
	Reverse	AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATACC	Pst I
138	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Pst I
141	Forward	AAAGAATTC-ATGAGCTTCAAACCCGATGCCGAAATCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Pst I
142	Forward	CGCGGATCCCATATG-CGTGCCGATTTTCATG	BamHI-
	Reverse	CCCGCTCGAG-AAACTGCTGCACATGGG	NdeI
143	Forward	AAAAAAGAATTC-ATGCTCAGTTTCGGCTTTCTCGGCGTTCAGAC	XhoI
	Reverse	AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTTCTTTAAT	Eco RI
144	Forward	AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Pst I
	Reverse	AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Eco RI
146	Forward	AAAAAAGAATTC-CGCCAAGTCGTCATTGACCACGACAAAGTC	Xba I
	Reverse	AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAACTGGG	Eco RI
147	Forward	AAAAAAGAATTC-ACTGAGCAATCGGTGGATTGGAAC	Pst I
	Reverse	AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG	Eco RI
148	Forward	AAAAAAGAATTC-ATGGCGTTAAAAACATCAAACCTTGGAACACGC	Xba I
	Reverse	AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTCG	BamHI-
149	Forward	CGCGGATCCCATATG-CTGCTTGACAACAAAGT	

	Reverse	CCCGCTCGAG-AAACTTCACGTTACGCC	NdeI
150	Forward	CGCGGATCCCATATG-CAGAACACAAATCCG	XhoI
			BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATAAACATCACGCTGATAGC	XhoI
151	Forward	AAAAAAGAATTC- ATGAAACAAATCCGCAACATCGCCATCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAGCTTTTAAAGTGGCGGCG	Pst I
152	Forward	AAAAAAGAATTC- ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	Pst I
153	Forward	AAAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTCATGTTTTCCGTTTCATT	Pst I
153a	Forward	AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCAGCATTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCGGCTTCCTTTCGGG	XhoI
155	Forward	AAAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCCCTTCTTAAACATATTCAGCAT	Pst I
156	Forward	AAAAAAGAATTC-GCACAGCAAAACGGTTTTGAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Pst I
157	Forward	CGCGGATCCCATATG-AGGAACGAGGAAAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAAACACAATATCCCCGC	XhoI
158	Forward	AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTATCCACAGAGATTGTTTCCCAGTTC	Xba I
160	Forward	CGCGGATCCCATATG-GACATTCTGGACAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	XhoI
163	Forward	AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	Kpn I
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
163a	Forward	AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Eco RI
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
164	Forward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACCTGC	XhoI
165	Forward	CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	BamHI-
			NheI
	Reverse	CCCGCTCGAG-AATATCCAATACTTTCGCG	XhoI
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	XhoI
209	Forward	CGCGGATCCCATATG-CTGCGGCATTTAGGA	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Forward	AAAAAAGAATTC-ATGTTGCGGGTTGCTGCTGC	Eco RI
	Reverse	AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward	CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGGGTAGATCCTTCC	XhoI
215	Forward	CGCGGATCCCATATG-GCATGGTTGGGTTCGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATATCTTTTGTATCATAAATC	XhoI
216	Forward	CGCGGATCCCATATG-GCAATGGCAGAAAACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward	CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACCCCGAATATCGAATCC	XhoI
218	Forward	CGCGGATCCCATATG-GTCGCGGTTCGATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TAACTCATAGAATCCTGC	XhoI
219	Forward	CGCGGATCCGCTAGC-ACGGCAAGGTTAAG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward	CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCTTCCCGCGTGTC	XhoI
225	Forward	CGCGGATCCCATATG-GACGAGTTGACCAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTTCAGAAAGCGGGAC	XhoI
226	Forward	AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	Eco RI
	Reverse	AAACTGCAG-TCAAAATCCCAAAACGGGGAT	Pst I
228	Forward	CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI
229	Forward	CGCGGATCCCATATG-CAAGAGGTTTGTCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACACAATATAGCGGATGAAC	XhoI
230	Forward	CGCGGATCCCATATG-CATCCGGGTGCCGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAGTTTGGCGGCTTCGG	XhoI
232	Forward	AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTGGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGGTTTTTCTGATTGCCGCCGC	Pst I
232a	Forward	AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCAGATG	Pst I
233	Forward	CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGCATTGAGCAG	XhoI
234	Forward	CGCGGATCCCATATG-GCCGTTTCACTGACCG	BamHI-

	Reverse	GCCCAAGCTT-ACGGTTGGATTGCCATG	NdeI
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	Hind III
	Reverse	CCCGCTCGAG-TTTGGGCTGCTCTTC	BamHI-
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	NdeI
	Reverse	CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	Forward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCCG	BamHI-
	Reverse	CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	NdeI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	XhoI
	Reverse	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	EcoRI-
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	NdeI
	Reverse	CCCGCTCGAG-AAACGCCATTACCCGATG	XhoI
241	Forward	CCGGAATTCTACATATG-CCAACACGTCCAAC	BamHI-
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	NdeI
242	Forward	CGCGGATCCCATATG-ATCGGCAAACCTTGTTG	XhoI
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	BamHI-
243	Forward	CGCGGATCCCATATG-ACGATTTTTTCGATGCTGC	NdeI
	Reverse	CCCGCTCGAG-CGACTTGTTACCGCG	HindIII
244	Forward	CGCGGATCCCATATG-CCGTCTGAAGCCC	BamHI-
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATT	NdeI
246	Forward	CGCGGATCCCATATG-GACATCGGCAGTGC	XhoI
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	BamHI-
247	Forward	CGCGGATCCCATATG-GTCGGATCGAGTTAC	NdeI
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGC	XhoI
248	Forward	CGCGGATCCCATATG-CGAAACAGAACACT	BamHI-
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	NdeI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	XhoI
	Reverse	CCCGCTCGAG-TTCCCGACCTCCGAC	BamHI-
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	NdeI
	Reverse	CCCGCTCGAG-TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253a	Forward	AAAAAAGAATTC-AAAATCCTTTTAAAAACAAGCGAAAACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward	AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
255	Forward	CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATCCGCAATACCGACCAG	XhoI
256	Forward	CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-ACGCCTGTTTGTGCGG	XhoI
257	Forward	CGCGGATCCCATATG-GCGGTTTCTTTCCTG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GCGCGTGAATATCGCG	XhoI
258	Forward	AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
258a	Forward	AAAAAAGAATTC-GCGAAGGCGGTGGCGCAAGGCCA	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
259	Forward	CGCGGATCCCATATG-GAAGAGCTGCCTCCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GGCTTTTCCGGCGTTT	XhoI
260	Forward	CGCGGATCCCATATG-GGTGCGGGTATGGT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AACAGGGCGACACCCT	XhoI
261	Forward	AAAAAAGAATTC-CAAGATACAGCTCGGGCATTTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	Pst I
263	Forward	CGCGGATCCCATATG-GCACGTTTAACCGTA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GGCGTAAGCCTGCAATT	XhoI
264	Forward	AAAAAAGGTACC-GCCGACGCAGTGGTCAAGGCAGAA	Kpn I
	Reverse	AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCC	Pst I
265	Forward	AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATACGTCGTCAAATG	Pst I
266	Forward	AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCCTGTTGCGCGTGTGCCA	Pst I
267	Forward	AAAGAATTC-TTCTTCCGATTTCGATGTTAATCG	Eco RI
	Reverse	AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward	AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward	AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
270	Forward	AAAGAATTC-GCCGTCAAGCTCGTTTGTGCAATG	Eco RI
	Reverse	AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Pst I
271	Forward	CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCCCAGCCCCGTGGAG	XhoI

272	Forward	AAAGAATTC-ATGACCGCAAAGGAAGAACTGTTTCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
273	Forward	AAAGAATTC-ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
	Reverse	AAACTGCAG-TTACGCGTAAGAAAAAACTGC	Pst I
274	Forward	CGCGGATCCCATATG-ACAGATTTGGTTACGGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCTTTCAGTATTATTGAA	XhoI
276	Forward	AAAAAAGAATTC- ATGATTTTGCCGTCGTCATCACGATGATGCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACACCACCATCGGCGAATTTATGGC	Pst I
277	Forward	AAAAAAGAATTC-ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
277a	Forward	AAAAAAGAATTC-GGGGCGGCGGCTGGGTTGGACGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
278	Forward	AAAAAAGGTACC-GTCAAAGTTGTATTAATCGGGCCTTTGCC	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
278a	Forward	AAAAAAGAATTC-AAAACTCTCCTAATTCGTCATAGTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA	XhoI
280	Forward	AAAAAAGGTACC-GCCCCCTGCCGTTTGTAAACCAG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTGCTTCATCGCGTTGGTCAAGGC	Pst I
281	Forward	AAAAAAGAATTC-GCACCCGTCGGCGTATTCCTCGTCATGCG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	AAAAAAGAATTC-TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282	Forward	AAAAAAGAATTC-CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283	Forward	CGCGGATCCCATATG-GCCGTCTTTACTTGGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGGCAGTATTTGTTTACG	XhoI
284	Forward	CGCGGATCCCATATG-TTGCCTGCAAAAGAATCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGACTTTGCAAAAACTG	XhoI
286	Forward	CGCGGATCCCATATG-GCCGACCTTTCCGAAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAAGCGCGTTCCCAAG	XhoI
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG	EcoRI- NheI
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTTGCC	XhoI
288	Forward	CGCGGATCCCATATG-CACACGGGACAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGTATCAAAGACTTGCGT	XhoI
290	Forward	CGCGGATCCCATATG-GCGGTTTGGGGCGGA	BamHI-

	Reverse	CCCGCTCGAG-TCGGCGCGGCGGGC	NdeI
292	Forward	CGCGGATCCCATATG-TGCGGGCAAACGCCC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-TTGATTTTTGCGGATGATTT	NdeI
294	Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC	XhoI
	Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Eco RI
295	Forward	CGCGGATCCCATATG-AACCGGCCGGCCTCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CGATATTTGATTCCGTTGC	NdeI
297	Forward	AAAAAAGAATTC-GCATACATTGCTTCGACAGAGAG	XhoI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Eco RI
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Pst I
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTACTTGATGGTTGCGT	Eco RI
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	NheI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	XhoI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGGATAGAATGTGGGGCGC	Eco RI
305	Forward	AAAAAAGGTACC-GAATTTTTACCGATTTCCAGCACCGGA	Pst I
	Reverse	AAAAAACTGCAG-TCATTCCCACTTATCCAGCCTGACAG	Kpn I
305a	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Pst I
	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Kpn I
306	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCGCATCGGCAGAC	NdeI
308	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	NdeI
311	Forward	AAAAAAGGTACC-ATGTTTCAGTTTTGGCTGGGTGTTT	XhoI
	Reverse	AAACTGCAG-ATGTTTCATATTCCTGCCTTCGGC	Kpn I
312	Forward	AAAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Pst I
	Reverse	AAACTGCAG-TCAGTTTTTTCATCGATTGAACCGG	Kpn I
313	Forward	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Pst I
	Reverse	AAAAAACTGCAG-TCAGCGGCTGCCGCCGATTTTGCT	Eco RI
401	Forward	CGCGGATCCCATATG-AAGGCGGCAACACAGC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	NdeI
402	Forward	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	XhoI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
402a	Forward	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Xba I
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	Xba I
			BamHI-
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG	NdeI
			XhoI

501	Forward	CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGTGTGATGTTACCC	XhoI
502	Forward	CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCTGCATGGCGGCG	XhoI
503	Forward	CGCGGATCCCATATG-TGTTGCGGGAAAGGCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGCGCATTCCTCGCA	XhoI
504	Forward	CGCGGATCCCATATG-AGCGATATTGAAGTGACG	BamHI- NdeI
	Reverse	GCCCAAGCTT-TGATTCAAGTCCTTGCCG	HindIII
505	Forward	CGCGGATCCCATATG-TTTCGTTTACAATTCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGTTTTATAGCGG	XhoI
510	Forward	CGCGGATCCCATATG-CCTTCGCGGACAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCACTGGCAGCG	XhoI
512	Forward	CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGAATAGCCTTTGACG	XhoI
515	Forward	CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATGCCGCAAAGCATC	XhoI
516	Forward	CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCGGGCGGCATC	XhoI
517	Forward	CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGCGCCAGCCGT	XhoI
518	Forward	AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	Eco RI
	Reverse	AAACTGCAG-TCAAATTCAGACTCTGCCAC	Pst I
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward	CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATATTACATTTCAAGTCGGC	XhoI
521	Forward	CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward	CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward	CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI- NdeI

Reverse	CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	XhoI
525 Forward	CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI- NdeI
Reverse	CCCGCTCGAG-GCCCGTGCATATCATAAA	XhoI
527 Forward	AAAGAATTC-TTCCCTCAATGTTGCCGTTTTTCG	Eco RI
Reverse	AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
529 Forward	CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI- NheI
Reverse	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
530 Forward	CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI- NdeI
Reverse	CCCGCTCGAG-ACGACCGACTGATTCCG	XhoI
531 Forward	AAAAAAGAATTC-TATGCCGCCGCTACCAAATCTACGG	Eco RI
Reverse	AAAAAACTGCAG-TTAAACAGCGCCGTGCCGACGACAAG	Pst I
532 Forward	AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Eco RI
Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Pst I
532a Forward	AAAAAAGAATTC-TTGGGTGTCGCGTTTGAGCCGGAAGT	Eco RI
Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Pst I
535 Forward	AAAGAATTC-ATGCCCTTTCCCGTTTTTCAGAC	Eco RI
Reverse	AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Pst I
537 Forward	CGCGGATCCCATATG-CATACCCAAAACCAATCC	BamHI- NdeI
Reverse	CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	XhoI
538 Forward	CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	BamHI- NdeI
Reverse	CCCGCTCGAG-TGGCATTTCGGTTTTCGTC	XhoI
539 Forward	CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	BamHI- NheI
Reverse	CCCGCTCGAG-TACCAATGTCGGCAAATC	XhoI
542 Forward	AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
Reverse	AAACTGCAG-TTACCGCGAACC GGTCAGGAT	Pst I
543 Forward	AAAAAAGAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
543a Forward	AAAAAAGAATTC-GGCAAAACTCGTCATGAATTTGC	Eco RI
Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
544 Forward	AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCGACCTGCACGG	Eco RI
Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
544a Forward	AAAAAAGAATTC- GCAATGACTATAAAAAACAAAAACTTCCAAGTACTTGC	Eco RI
Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
547 Forward	AAAGAATTC-ATGTTCTGTAGATAACGGATTTAATAAAAC	Eco RI
Reverse	AAACTGCAG-TTAACAACAAAAACAAACCGCTT	Pst I
548 Forward	AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI

	Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	Pst I
550	Forward	AAAAAAGTCGAC- ATGATAACGGACAGGTTTCATCTCTTTTCATTTTCC	Sal I
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
550a	Forward	AAAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
552	Forward	AAAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	Eco RI
	Reverse	AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Pst I
554	Forward	CGCGGATCCCATATG-TCGCCCGCGCCCAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	XhoI
556	Forward	AAAGAATTC-GCGGGCGGTTTTGTTTGGACATCCCG	Eco RI
	Reverse	AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Pst I
557	Forward	CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	XhoI
558	Forward	AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	Eco RI
	Reverse	AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
558a	Forward	AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
560	Forward	AAAAAAGAATTC- TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGCGGTTTTCAGACGGCATTTTGGC	Pst I
561	Forward	CCGGAATTCTACATATG-ATACTGCCAGCCCGT	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTCAAGCTTTCTTCAGATG	XhoI
562	Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGACCAACTCCAACCTCGT	XhoI
565	Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566	Forward	CGCGGATCCCATATG-GTCGGTGGCGAAGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567	Forward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTCCCCGACACCCTCG	XhoI
568	Forward	CGCGGATCCCATATG-CTCAGGGTCAGACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGCGGCGTTCAG	XhoI
569	Forward	AAAAAAGAATTC-CTGATTGCCTTGTGGGAATATGCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Pst I
570	Forward	CGCGGATCCCATATG-GACACCTTCAAAAAAATCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGGCGTTCATTTCTTT	XhoI

571	Forward	AAAAAAGAATTC- ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATGGCCGACGCGCGGCTACCTGACG	Pst I
572	Forward	CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCAGTGTGCCGATA	XhoI
573	Forward	CGCGGATCCCATATG-CCCTGTTTGTGCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGTGTCAATTCGCC	XhoI
574	Forward	CGCGGATCCCATATG-TGTTTTGCCGCCCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACTTCGATTTTATTCGGG	XhoI
575	Forward	CGCGGATCCCATATG-GTTTCGGGCGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCGAATCTGAACAG	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	XhoI
577	Forward	CGCGGATCCCATATG-GAAAGGAACGGTGTATTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGCTGTTTGGTAGATTCTG	XhoI
578	Forward	CGCGGATCCCATATG-AGAAGGTTTCGTACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward	CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
580	Forward	CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACTTCCCCCGAAGTG	XhoI
581	Forward	CGCGGATCCCATATG-CACTTCGCCCAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCCGTTTGGCTTTGG	XhoI
582	Forward	AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATGCCGTCCAGTCGTTGAA	Xba I
583	Forward	AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Pst I
584	Forward	AAAAAAGAATTC- GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGAACTGAACCGTCCCATTGACGCT	Pst I
585	Forward	AAAAAAGGTACC-TCTTCTGGCTGGTGCAGAACACCCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTTCGCACTTTTTCTGTTTGGGA	Pst I
586	Forward	CGCGGATCCCATATG-GCAGCCCATCTCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	XhoI
587	Forward	CGCGGATCCCATATG-GACCTGCCCTTGACGA	BamHI-

	Reverse	CCCGCTCGAG-AAATGTATGCTGTACGCC	NdeI
			XhoI
588	Forward	AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC- ATGCAACAAAAAATCCGTTTCCAAATCGAAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAATCGATTTTACCCGTTTCAGGCG	Pst I
590	Forward	AAAAAAGAATTC-ATGAAAAAACCTTTGATTTCAAGTTGCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591	Forward	AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAACCGAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAAACGTCGGATTTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACCTGAACGGACTC	BamHI- NheI
	Reverse	CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a	Forward	AAAAAAGAATTC-GGTAAGTTCGCCGTTTCAGGCCTTTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGCCGTTTCTTGACACTCGCG	Pst I
595	Forward	AAAAAAGAATTC-TGCCAGCCGCCGGAGGCGGAGAAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
597	Forward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI- NdeI
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	HindIII
601	Forward	CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	XhoI
602	Forward	CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI- NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCTCGCGTAGGCGGGGACGGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604	Forward	CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI- NheI
	Reverse	CCCGCTCGAG-GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAGCGATTTGAGGCGGGCGATACG	Pst I
607	Forward	AAAAAAGAATTC-ATGCTGCTCGACCTCAACCGCTTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608	Forward	AAAAAAGAATTC-ATGTCCGCCCTCCTCCCCATCATCAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTCTATCCAAATGTCGCGTTC	Pst I
609	Forward	CGCGGATCCCATATG-GTTGTGGATAGACTCG	BamHI- NdeI

SUBSTITUTE SHEET (RULE 26)

SUBSTITUTE SHEET (RULE 26)

Reverse	CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
610 Forward	CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI- NdeI
Reverse	CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611 Forward	CGCGGATCCCATATG-CCGTCTCAAAACGGG	BamHI- NdeI
Reverse	CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
613 Forward	CGCGGATCCCATATG-TCGCGTTCGAGCCG3	BamHI- NdeI
Reverse	CCCGCTCGAG-AGCCTGTAAAATAAGCGGC	XhoI
614 Forward	CGCGGATCCCATATG-TCCGTCGTGAGCGGC	BamHI- NdeI
Reverse	CCCGCTCGAG-CCATACTGCGGCGTTC	XhoI
616 Forward	AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTCGG	Eco RI
Reverse	AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
619a Forward	AAAAAAGAATTC-GGGCTTCTCGCCGCTCGCTTGC	Eco RI
Reverse	AAAAAATCTGCAG-TCATTTTTTGTGTTTTAAACGAGATA	Pst I
622 Forward	CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI- NdeI
Reverse	CCCGCTCGAG-TTTGTCCAAATGATAAATCTG	XhoI
624 Forward	CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI- NdeI
Reverse	CCCGCTCGAG-AGATTCGGGCCTGCGC	XhoI
625 Forward	CGCGGATCCCATATG-TTTGCAACCAGGAAAATG	BamHI- NdeI
Reverse	CCCGCTCGAG-CGGCAAAATTACCGCCTT	XhoI
627a Forward	AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	Eco RI
Reverse	AAAAAACTGCAG-TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628 Forward	AAAAAAGGTACC-GCCTTACAAACATGGATTTTGCGTTC	Kpn I
Reverse	AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAAA	Pst I
629a Forward	AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Eco RI
Reverse	AAAAAACTGCAG-TTACAACACCGCCGTCGGTTCAAACC	Pst I
630a Forward	AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTCCG	Eco RI
Reverse	AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Pst I
635 Forward	AAAAAAGAATTC-ATGACCCAGCGACGGGTCTGGCAAGCAAAACCG	Eco RI
Reverse	AAAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	Pst I
638 Forward	AAAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Eco RI
Reverse	AAAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Pst I
639 Forward	CGCGGATCCCATATG-ATGCTTTATTTTGTTCG	BamHI- NdeI
Reverse	CCCGCTCGAG-ATCGCGGCTGCCGAC	XhoI
642 Forward	CGCGGATCCCATATG-CGGTATCCGCCGCAAT	BamHI- NdeI
Reverse	CCCGCTCGAG-AGGATTGCGGGGCATTA	XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAACCGAAAAACAGACCGC	XhoI
644	Forward	AAAAAAGAATTC- ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	Eco RI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645	Forward	AAAAAAGAATTC-GTGAACAGAGCAACACGTAAATCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647	Forward	AAAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648	Forward	AAAAAAGAATTC- ATGAACAGGCGCGACGCGCGGATCGAACG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649	Forward	AAAAAAGAATTC-GGTACGTGAGAACCCGCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAACTGCCGCCGTC	Pst I
650	Forward	AAAAAAGAATTC-ATGTCCAACTCAAACCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652	Forward	AAAAAAGGTACC- GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTTGCCAGTTGGTAGAATGCGGC	Pst I
653	Forward	AAAAAAGAATTC-GCGGCTTTGCCGTAATTTTCATCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a	Forward	AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-CTACGATTTGCGCGATTTCCACATCGT	Pst I
657	Forward	AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst I
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC- ATGCACATCGGCGGCTATTTTATCGACAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCACGACGTGTCTGTTCCGCCGTCGGGC	Pst I
663	Forward	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	XhoI
664	Forward	CGCGGATCCCATATG-GCGGCTGGCGCGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665	Forward	AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAAAATTTTGCCGACGATTTTC	Pst I
666	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667	Forward	AAAAAAGAATTC- CCGCATCCGTTTGATTTCCATTTTCGTATTCGTCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward	AAAAAAGAATTC-ATGCGCCGCATCATTAACACACCAGCC	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTATCCGTTTGATGTCGGC	Pst I
670a	Forward	AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTCGTCTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGAGCTTTTGGAACGCGTCGGACTGGC	Pst I
671	Forward	CGCGGATCCCATATG-ACCAGCAGGGTAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAACTATAAAAACGCAAG	XhoI
672	Forward	CGCGGATCCCATATG-AGGAAAATCCGCACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGGGATAGGCGGTTG	XhoI
673	Forward	AAAAAAGAATTC-ATGGATATTGAAACCTTCCTTGCAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA	Pst I
674	Forward	AAAAAAGAATTC-ATGAAAACAGCCCGCCGCGTTCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG	Pst I
675	Forward	CGCGGATCCCATATG-AACACCATCGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCTTCGTCTTCAAACCTGT	XhoI
677a	Forward	AAAAAAGAATTC-AGACGGCATTCGCGATCAGTCGATTTTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC	Pst I
680	Forward	CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG	BamHI- NheI
	Reverse	CCCGCTCGAG-CATCAAAAACCTGCCGC	XhoI
681	Forward	AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCGTCTTCCGCAAAAAACAGC	Pst I
683	Forward	CGCGGATCCCATATG-TGCAGCACACCGGACAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAGTTTTTTTCCGCATACG	XhoI
684	Forward	CGCGGATCCCATATG-TGCGGTACTGTGCAAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTCGACCATCTGTTGCG	XhoI
685	Forward	CGCGGATCCCATATG-TGTTTGCTTAATAAATAACATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTTTTCCCGCCGCA	XhoI
686	Forward	CGCGGATCCCATATG-TGCGGCGGTTTCGGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGATTCTGATGAAG	XhoI
687	Forward	CGCGGATCCCATATG-TGCGACAGCAAAGTCCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCGCGGCTTTTTGTT	XhoI
690	Forward	CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TATTCGCCCCGTGTTTGG	XhoI
691	Forward	CGCGGATCCCATATG-GCCACGGCTTATATCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGGCAGGAAGAAAG	XhoI

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCTGCGTCGGTGCGGT	XhoI
695	Forward	CGCGGATCCCATATG-TTGCCTCAAACCTCGTCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
696	Forward	CGCGGATCCCATATG-TTGGGTGCGCGCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	XhoI
700a	Forward	AAAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
701	Forward	CGCGGATCCCATATG-AAGACTTGTGTTGGATACTTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702	Forward	AAAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACCCCATTCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a	Forward	AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTTTTGCCGGATAATATGGCGGGTGCG	Pst I
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI- NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TTGACCGGTGAGGACG	XhoI
710	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
714	Forward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
716	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	Forward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGCGGGCATGGTCTTGTC	XhoI
720	Forward	CGCGGATCCCATATG-AGCGGATGGCATAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	XhoI
723	Forward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	XhoI

725	Forward	CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	XhoI
726	Forward	CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCGATGTTTAGCGTCC	XhoI
728	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
729	Forward	CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTTGTGCGTTTGGGTATC	HindIII
731	Forward	CGCGGATCCGCTAGC-GCCGTGCCGGAGG	BamHI- NheI
	Reverse	CCCGCTCGAG-ACGGGCGCGGCAG	XhoI
732	Forward	CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	EcoRI- NdeI
	Reverse	CCCGCTCGAG-CTTCTTATCTTTTTTATCTTTC	XhoI
733	Forward	CGCGGATCCCATATG-GCCTGCGGCGGCAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGCTTGCCCTCCTTTAC	XhoI
734	Forward	CGCGGATCCCATATG-GCCGATACTTACGGCTAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	XhoI
735	Forward	CGCGGATCCCATATG-AAGCAGCAGGCGGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
737	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	XhoI
739	Forward	CGCGGATCCCATATG-GCAAAAAAACC GAACA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAAGAGTTTGTGCGAGAATT	XhoI
740	Forward	CGCGGATCCCATATG-GCCAATCCGCCC GAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
741	Forward	CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTGCGGCAAGGC	XhoI
743	Forward	CGCGGATCCCATATG-GACGGTGTTGTGCCTGTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTACGGATCAAATTGACG	XhoI
745	Forward	CGCGGATCCCATATG-TTTTGGCAACTGACCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
746	Forward	CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-

	Reverse	CCCGCTCGAG-TTCATTCGTTACCTGACC	NdeI
747	Forward	CGCGAATTCTAGCTAGC-CTGACCCCTTGGG	XhoI
	Reverse	GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC	EcoRI-
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG	NheI
	Reverse	CCCGCTCGAG-TTTC AAGCCGAGTATGC	HindIII
750	Forward	CGCGGATCCCATATG-TGTTGCCCCGAACCTG	BamHI-
	Reverse	CCCGCTCGAG-CTTTTCCCCGCCGCAA	NdeI
758	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT	XhoI
	Reverse	CCCGCTCGAG-TGGCTCAATCCTTTCTGC	BamHI-
759	Forward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC	NdeI
	Reverse	CCCGCTCGAG-CCAGTTGTAGCCTATTTTG	XhoI
763	Forward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG	BamHI-
	Reverse	CCCGCTCGAG-TTCCGCAAATACCGTTTCC	NdeI
764	Forward	CGCGGATCCCATATG-TTTTCTCCGCCCTGA	XhoI
	Reverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC	BamHI-
765	Forward	CGCGGATCCCATATG-TTAAGATGCCGTCCG	NdeI
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTTATTAA	XhoI
767	Forward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG	NdeI
768	Forward	CGCGGATCCCATATG-GCCCCGCAAAAACCCG	XhoI
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC	BamHI-
770	Forward	CGCGGATCCCATATG-TGCGGCAGCGGCGAA	NdeI
	Reverse	CCCGCTCGAG-GCGTTTGTGCGAGATTTTC	XhoI
771	Forward	CGCGGATCCCATATG-TCCGTATATCGCACCTTC	BamHI-
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG	NdeI
772	Forward	CGCGGATCCCATATG-TTTCGGCGTTGGTGG	XhoI
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG	BamHI-
774	Forward	CGCGGATCCCATATG-TCCGTTTCACCCGTTCC	NdeI
	Reverse	CCCGCTCGAG-TCGTTTGC GCACGGCT	XhoI
790	Forward	CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC	BamHI-
			NdeI

Reverse	CCCGCTCGAG-GGCGTTGTTCCGATTTCG	XhoI
900 Forward	CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI- NdeI
Reverse	CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC	XhoI
901 Forward	CGCGGATCCCATATG-CCCATTTCGATG	BamHI- NdeI
Reverse	CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
902 Forward	CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC	EcoRI- NdeI
2		
Reverse	CCCGCTCGAG-AAAAATGTACAATGGCGTAC	XhoI
903 Forward	CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT	EcoRI- NheI
Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA	XhoI
904 Forward	AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
904a Forward	AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
905 Forward	CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI- NdeI
Reverse	CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
907 Forward	CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI- NdeI
Reverse	CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
908 Forward	AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
Reverse	AAACTGCAG-TTAATATGGTTTTGTCGTTTCG	Pst I
909 Forward	CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI- NdeI
Reverse	CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT	XhoI
910 Forward	AAAGAATTC-GCATTGCGCGGCGACTCTGCCGAGCG	Eco RI
Reverse	AAACTGCAG-TCAGCGATCGAGCTGCTCTTT	Pst I
911 Forward	AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC	Eco RI
Reverse	AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTCCGC	Pst I
912 Forward	AAAAAAGAATTC- CAAATCCGTCAAAACGCCACTCAAGTATTGAG	Eco RI
Reverse	AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC	Pst I
913 Forward	CGCGGATCCCATATG-GAAACCCGCCCCGC	BamHI- NdeI
Reverse	CCCGCTCGAG-AGGTTGTGTTCCAGGTTG	XhoI
915 Forward	CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI- NdeI
Reverse	CCCGCTCGAG-TTTGAAAATATAGGTATCAGG	XhoI
914 Forward	AAAGAATTC-GACAGAATCGGCGATTGGAAGCACG	Eco RI
Reverse	AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGG	Pst I
916 Forward	CGCGGATCCCATATG-GCAATGATGGCGGCTG	BamHI- NdeI
Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI

917	Forward	AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG	NdeI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	XhoI
	Reverse	CCCGCTCGAG-ATGGTGCGAATGACCGA	BamHI-
921	Forward	AAAAAAGAATTC-TTGACGGAAATCCCCGTGAATCC	NdeI
	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	XhoI
922	Forward	CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	Eco RI
2	Reverse	CCCGCTCGAG-CAATCCCGGGCCGCC	Pst I
923	Forward	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	NheI
925	Forward	CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	XhoI
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTT	BamHI-
926	Forward	CGCGGATCCCATATG-TGCGCGCAATTACCTC	NdeI
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	XhoI
927	Forward	CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-
	Reverse	CCCGCTCGAG-GTTTTTGTCTGACGTAGT	NdeI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	XhoI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAACTACTGCC	Eco RI
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Xba I
	Reverse	AAAAAACTGCAG-TTACTGCCCGACAACAACGCGACG	Eco RI
935	Forward	AAAAAAGAATTC-GCGGATGCGCCCGCGATTTTGGATGACAAGGC	Pst I
	Reverse	AAAAAACTGCAG-TCAAAACCGCCAATCCGCCGACAC	Eco RI
936	Forward	CGCGGATCCCATATG-GCCGCCGTGCGCGC	Pst I
	Reverse	CCCGCTCGAG-GCGTTGGACGTAGTTTTG	BamHI-
937	Forward	AAAAAAGAATTC-CCGTTTACATTCAAACCGGCGCAAC	NdeI
	Reverse	AAAAAACTGCAG-TTAAATGTATGCTGTACGCCAAA	XhoI
939a	Forward	AAAAAAGAATTC-GGTTCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	Eco RI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	BamHI-
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	NdeI
	Reverse	CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	XhoI
957	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-

	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	NdeI
958	Forward	CGCGGATCCCATATG-GCCGATGCCGTTGCG	XhoI
			BamHI-
	Reverse	GCCCAAGCTT-GGGTCGTTTGTTCGCTC	NdeI
959	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	HindIII
			BamHI-
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	NdeI
961	Forward	CGCGGATCCCATATG-GCCACAAGCGACGACG	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-CCACTCGTAATTGACGC	NdeI
972	Forward	AAAAAAGAATTC- TTGACTAACAGGGGGGGAGCGAAATTA AAAAC	XhoI
			Eco RI
	Reverse	AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	Xba I
973	Forward	AAAAAAGAATTC-ATGGACGGCGCACAACCGAAAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTTCACGCGGGTCGCCATCAGCGT	Pst I
982	Forward	CGCGGATCCCATATG-GCAGCAAAGACGTAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CATCATGCCGCCCATCC	XhoI
983	Forward	CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
987	Forward	CGCGGATCCCATATG-CCCCACTGGAAGAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TAATAAACCTTCTATGGGC	XhoI
988	Forward	CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	BamHI-
			NdeI
	Reverse	GCCCAAGCTT-TGATTTGCCTTTCCGTTTT	HindIII
989	Forward	CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	EcoRI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	XhoI
990	Forward.	CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	BamHI-
	2		NheI
	Reverse	CCCGCTCGAG-AAACAGCCATTTGAGCGA	XhoI
992	Forward	CGCGGATCCCATATG-GACGCGCCCGCCCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCAAATGCCCAACCATTC	XhoI
993	Forward	CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAACACATCGCGCCCG	XhoI
996	Forward	CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCTAAACCCCTGTTTTCTC	XhoI
997	Forward	CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	EcoRI-
			NheI
	Reverse	CCCGCTCGAG-GACGGCATCGCTCAGG	XhoI

Underlined sequences indicate restriction recognition sites.

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 1>:

```
g001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG GTGTCGGCGA ACGAGGTGTC
51  CGGCAGGGCT TGCGCCCGGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
101 AACGCGATAC TTTAAACGGC TCGGGTACGC ATACTTTACC GGTTCGGCG
151 ATTTTGCCGA GGTCGTTGCG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCGGTTT GTAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTTCGGT GCTGATGGTG
301 CCGTCTGAAG CGATGTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCGGCTTCAT CGGGCAGGTG GGACAATACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:

```
g001.pep
1  MLPQGKAARR VSANEVSGRA CARMVLVICQ TLPKRDTLNG SGTHTLPVWA
51  ILPRSLRSKS TIITFSARFF GSVNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEAMLRKSS GEKHSVHAD C PASSGRWDNT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3>:

```
m001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
51  CGGCAssCTT ss.GCTTGGGA YGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTCGGCG
151 ATTTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTTCGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCCTCCGCAT CGGGCAGGTG GGACAAGACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:

```
m001.pep
1  MLPQGKAARR MSANEVCGXL XAWXVLVICQ TLPKRDTLNG SGTHTVPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C PSASGRWDKT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 5>:

```
a001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
51  CGGCAAGGCT TGGGCTTGGGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTCGGCG
151 ATTTTGCCGA GGTCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTTCGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
351 CGCGGATTGC CCTGTGTCAT CGGGCAGGTG GGACAAAACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:

```
a001.pep
1  MLPQGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGTHTVPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C PCASGRWDKT A*
```

m001/a001 96.2% identity over a 131 aa overlap

```

10      20      30      40      50      60
m001.pep  MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
          |||
a001.pep  MLPQGKAARRMSANEVCGKAWAWMVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
```

166

	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHAD					
a001.pep	TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHAD					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
a001.pep	PCASGRWDKTAX					
	130					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 001 shows 89.3% identity over a 131 aa overlap with a predicted ORF (ORF 001.ng) from *N. gonorrhoeae*:

m001/g001

	10	20	30	40	50	60
m001.pep	MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS					
g001	MLPQGKAARRVSANEVSGRACARMVLVICQTLPKRDTLNGSGTHTLPVWAILPRSLRSKS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHAD					
g001	TIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHAD					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
g001	PASSGRWDNTAX					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 7>:

g003.seq

```

1  ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTGCGCTG TTTTGGGTCA
51  CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCAGTG CGGCTTTGGT
151 TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCCTCGTAG
251 AAGTTTTTCA GCGGTTTCGT TTCAACAACG AGGGGCAGTT GGTATTTTGT
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351 TGAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451 CAGgCTGagT TTGTCGGCAT CGTaggtcac tTTGACGGTC TCGGCATGAC
501 CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTGttGCCG
551 TTGGCgttac cGGATACCGC gtcaACCACG CCGTcgatgc gttggaAATa
601 ggCTTCCAAG ccccaaaagc agccgccggc gaagtaaagt gtgcccgtgt
651 tcatgattGC TGa

```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>:

g003.pep

```

1  MVVFVAEGVF GRAVLGHLVL LFGQGAFFFG VTRFFIRCRV EAFALRCGFG
51  FARQRFVGFA DVDVAVAVGV FNQVVLNVFL GVVEVFQRFV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGVVH AAIVLRAGVV TLFVEAGRIN DAEIILQDVV

```

151 QAEFVGIVGH FDGLGMTRMA VGHEFVRVFR VAVGVTGYRV NHAVDAL EIG
201 FQAPKAAAGE VNGARVHDC

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 9>:

m003.seq
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCGCGCTG TTTTGGGTAA
51 CTTGsTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACCTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGGG CCGTCTTGGT
151 TTTGCCCGGC AGCGGTTCGT CAGCKTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251 AAGTTTTtCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGk CGGCGATGAC GGCTTTTTCG kCGGGGTCGG
351 TGTA GTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCGTT GCCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGCG
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATAA
601 GGCTTCCAAG CCCCAGAAGC AGCg.CCGGC GAGGTAAATG GTGCGCGTGT
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

m003.pep Length: 221
1 MVVFVAEGIF GRAVLGNLXL LFGQGAFFEG VTRFFIRCRV EAFALRGGLG
51 FARQRFVSHA DDDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL
101 LLAFEGXGDD GFFXGVGVVH AAVLRTGVV ALFVEAGRIN DAEEILQDVV
151 *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDAL E
201 GFQAPEAAXG EVNGARVHDF *

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 11>:

a003.seq
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCGCGCTG TTTTGGGTAA
51 CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACCTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CCGTCTTGGT
151 TTTGCCCGGC AGCGGTTCGT CCGCTTTGCG GATATCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251 AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351 TGTA GTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCGTT GCCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGCG
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATAA
601 GGCTTCCAAG CCCCAGAAGC AGCCGCCCGC GAGGTAGATG GTGCGCGTGT
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>:

a003.pep
1 MVVFVAEGIF GRAVLGNLVL LFGQGAFFEG VTRFFIRCRV EAFALRCGLG
51 FARQRFVGFA DIDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGVVH AAVLRTGVV ALFVEAGRIN DAEEILQDVV
151 *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDAL E
201 GFQAPEAAAAG EVDGARVHDF *

m003/a003 95.9% identity over a 220 aa overlap

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXL	LFGQGAFFEGVTRFFIRCRV	EAFALRGGLGFARQRFVSHA			
a003	MVVFVAEGIFGRAVLGNLVL	LFGQGAFFEGVTRFFIRCRV	EAFALRCGLGFARQRFVGFA			
	10	20	30	40	50	60

168

	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVLMLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH					
a003	DIDVAVAVGVFNQVVLMLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH					
	70	80	90	100	110	120
	130	140	150	160	170	180
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
a003	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
	130	140	150	160	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAVDALEIGFQAPEAAXGEVNGARVHDFX					
a003	RVAVGVAGYRVNHAVDALEIGFQAPEAAAGEVDGARVHDFX					
	190	200	210	220		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from *N. gonorrhoeae*:

m003/g003

	10	20	30	40	50	60
m003.pep	MVFVFAEGIFGRAVLGNLXLLFGQGAFFEGVTRFFIRCRVEAFALRGGLGFARQRFVSXA					
g003	MVFVFAEGVFGRAVLGHLVLLFGQGAFFEGVTRFFIRCRVEAFALRCGFGFARQRFVGFA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVLMLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH					
g003	DVDVAVAVGVFNQVVLMLVFLGVVEVFQRFVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH					
	70	80	90	100	110	120
	130	140	150	160	170	180
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
g003	AAAVLRAGVVTFLFVEAGRINDAEEILQDVVQAEFVGIVGHFDGLGMTRMAVGHFFV-RVF					
	130	140	150	160	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAVDALEIGFQAPEAAXGEVNGARVHDFX					
g003	RVAVGVTGYRVNHAVDALEIGFQAPKAAAGEVNGARVHDC					
	190	200	210			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 13>:

g004.seq

1	ATGgtagAAC	GGCATATCCA	GCATTTGCGG	AACGGTCATC	TTCATTTGAT
51	GCGCCCATGC	CAACAagtg	gccAAAtgtT	CGGCGGCAGG	GCCTacgatT
101	TCCGCGCCGA	TAAagcgcc	gGTGgctTTT	tcgGCataca	ggcgcaTatg
151	gCCTTTGTTT	ACCAgcatca	cgcggtcg	accttgATTT	TTGAACGATA
201	CTTCGCCgaT	GACAAATTCG	TCCGCTTGGT	ATTGCGCGGC	AACCTGCGCG
251	TATTTCAAAC	CGACAAAGCC	GATTTGCgga	ctggtaaACA	CCACGCCAAT
301	GGTgctgcgg	cGCAAACCGC	TGCCGATATt	cgGgtagcgg	ccccgcgtta
351	ttgcccggca	atcttacctt	ggtcggcggc	ttcatGCAGC	AGGGGCagtt
401	ggttggaacg	gtcgcccgca	ataAAGATAT	GCGGAATgct	ggtCTGCATg
451	gtCAGCGGAT	CGGCAACGGG	tacgcccgcg	gcgtctttgT	CGATATTGAT
501	GTTTTCCAAA	CCGATATgtT	CAACGTTCGG	ACGGCgACCT	ACGGCTGCCA

```

551 ACATATATTC GGCAACAAAT ACGCCTTTTT CGCCATCCTG CTCCCAATGG
601 ACTtctACAT TGCCGTCTGC GTCGAGTTTG ACCTCGGTTT TAGCATCCAG
651 ATGCAGTTTC AATtctTCTC CGAACACGGC TTTCGCCTCG TCTGAAACAA
701 CGGGGTCGGA AATGCCGCCG ATGATTCCGC CCAAACCGAA AATTTCAACT
751 TTCACACCCA AACGGTGCAA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>:

g004 . pep

```

1 MVERHIQHLR NGHLLHMRPC QQVSQMFGR AYDFRADKAA GGFFGIQAHM
51 AFVYQHHAHA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
101 GAAAQTAADI RVAAPRYCPA ILPWSAASCS RGSWLDASPA IKICGMLVCM
151 VSGSATGTFR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
201 TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMP MIPPKPKIST
251 FTPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 15>:

m004 . seq

```

1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
51 GTGCCCAGC CAACAGGTGC GCCAAATGTT CGGCGGCAGG GCCTACGATT
101 TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTC ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTGCGCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTACAGC CGACAAAGCC GATTTGCGGA CTGGTAAACA CCACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT
401 TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC
451 AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATATT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTTAG CATCCAGATG
651 CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATTCGCCCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

m004 . pep

```

1 MVERHIQHLR NGHLLHMCPS QQVRQMFGR AYDFRADKAA GGFFGIQAHM
51 AFVHQHHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101 GAAPQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAM KICGILVCMW
151 SGSATGTPRA SFSILIFSKP ILSTFGRRPT AASIYSATNT PFSPPSCSQW
201 STLPSASSLT SVLASRCSFN SSPNTAFASS ETTGSEMPPM IPPPKPKISTF
251 TPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 17>:

a004 . seq

```

1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
51 GTGCCCAGC CAACAGGTGC GCCAAATGTT CGGCGGCCG ACCTACGATT
101 TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTGCGCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTTGCGGA CTGGTGAACA CTACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT
401 TGGACGCGTC GCCCGCAATA AAGATATGCG GAATACTGGT CTGCATAGTC
451 AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATGTT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCTACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCGTC GAGTTTGCC TCGGTTTTAG CATCCAAATG
651 CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATGCCACCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```


This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>:

a004.pep

```

1  MVERHIQHLR NGHLHLMCPS QQVRQMFGGR TYDFCADEAA GGFFGIQAHM
51 AFVYQHAAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD
101 GAAAQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV
151 SGSATGTPRA SFSILMFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF
251 TPKRCNA*

```

m004/a004 94.9% identity over a 257 aa overlap

	10	20	30	40	50	60
m004.pep	MVERHIQHLR	NGHLHLMCPS	QQVRQMFGGR	TYDFRADKA	AAGGFFGIQAHM	AFVHQHHAAA
a004	MVERHIQHLR	NGHLHLMCPS	QQVRQMFGGR	TYDFRADKA	AAGGFFGIQAHM	AFVHQHHAAA
	70	80	90	100	110	120
m004.pep	ALVFERYFAD	DKFVGLVLRG	NLRVFQTDK	ADLRTGKHH	ADGAAPQTA	ADIRVAAALSPAI
a004	ALVFERYFAD	DKFVGLVLRG	NLRVFQTDK	ADLRTGKHH	ADGAAPQTA	ADIRVAAALSPAI
	130	140	150	160	170	180
m004.pep	LPWSAASCSR	GSWLDASPA	MKICGILVCM	VSATGTPRAS	SFSILIFSKP	ILSTFGRRPT
a004	LPWSAASCSR	GSWLDASPA	MKICGILVCM	VSATGTPRAS	SFSILIFSKP	ILSTFGRRPT
	190	200	210	220	230	240
m004.pep	AASIYSATNT	PFSPSCSQWT	STLPSASSL	TSVLASRCS	FNSSPNTAF	ASSETTGSEMPPM
a004	AASIYSATNT	PFSPSCSQWT	STLPSASSL	TSVLASRCS	FNSSPNTAF	ASSETTGSEMPPM
	250					
m004.pep	IPPKPKISTF	TPKRCNAX				
a004	MPPKPKISTF	TPKRCNAX				

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from *N. gonorrhoeae*:

m004/g004

	10	20	30	40	50	60
m004.pep	MVERHIQHLR	NGHLHLMCPS	QQVRQMFGGR	TYDFRADKA	AAGGFFGIQAHM	AFVHQHHAAA
g004	MVERHIQHLR	NGHLHLMRCP	QQVSQMFGGR	TYDFRADKA	AAGGFFGIQAHM	AFVHQHHAAA
	70	80	90	100	110	119
m004.pep	ALVFERYFAD	DKFVGLVLRG	NLRVFQTDK	ADLRTGKHH	ADGAAPQTA	ADIRVAAA-LSPA
g004	TLIFERYFAD	DKFVGLVLRG	NLRVFQTDK	ADLRTGKHH	ANGAAAQTA	ADIRVAAPRYCPA
	120	130	140	150	160	179
m004.pep	ILPWSAASCSR	GSWLDASPA	MKICGILVCM	VSATGTPRAS	SFSILIFSKP	ILSTFGRRP

```

g004      |||||:|||||:|||||:|||||:|||||:|||||
          ILPWSAASCSRGSWLDASPAIKICGMLVCMVSGSATGTPRASLSILMFSPILSTFGRRP
          130      140      150      160      170      180

m004.pep  180      190      200      210      220      230      239
          TAASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
g004      |||||:|||||:|||||:|||||:|||||:|||||
          TAANIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
          190      200      210      220      230      240

m004.pep  240      250
          MIPPKPKISTFTPKRCNAX
          |||||
g004      MIPPKPKISTFTPKRCNA
          250

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 19>:

```

g005.seq
1  ATGGGGATGG ACAATATTGA TATGTTTCATG CCTGAACAAG AGGAAATCCA
51  ATCAATGTGG AAAGAAATTT TACTGAATTA CCGTATTTTC CTGCTCGAAC
101 TGCTTACCGT GTTCGGCGCA ATTGCGCTGA TTGTGTTGGC TATCGTACAG
151 AGTAAGAAAC AGTCGGAAG CGGCAGTGTC GTACTGACAG ATTTTTCGGA
201 AAATTATAAA AACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
251 AAGAGACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGGCGGAA
301 GCCAAAGCAG AGAAAAAGCG TTTGAAGGAG GGCGGGGAGA AATCTGCCGA
351 AACGCAAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTGTATG
401 CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
451 GCCAAGCCTG AAGATGAGGT TCTGCTCAGA TTGGAAAGTC CGGGCGGCGT
501 GGTTACCGGT TACGGTTTGG CGGCTTCGCA GCTTAGGCGT TTGCGCGAAC
551 GCAATATTCC GCTGAccgtc gccgTCGATA AGGTCGCGGC AAGCGgcggc
601 tatatgatgg cgtgtgtgGC GGATAAAATT GTTTCGCTc cgtttgcggt
651 catcggttcg gtgggtgtgg tgGcggaagt gcCGAATATC CAccgCctGT
701 TGAAAAACA TGATATTGAT GTGGATGTGA TGACGGCGGG CGAATTTAAG
751 CGCACGGTTA CTTTATAGGG TGAAATACG GAAAAGGGCA AACAGAAATT
801 CCGGCAGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCAAGT
851 AAAACCGCCC CGGGTTGGAT ATTGAAAAAA TAGCGACGGG CGAGCATTGG
901 TTCGGCCGGC AGGCGTTGGC GTTGAAGTTG ATTGACGAGA TTTGACACAG
951 TGATGATTG TTGTTGAAAG CGTTTGAAAA CAAACAGGtt aTCGAAGTGA
1001 AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
1051 GCTTCGGTTG AAAAGTTGTT TGCCAACTT GTCAACCGGC GAGCGGATGT
1101 GATGTAG

```

This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

```

g005.pep
1  MGMDNIDMFEM PEQEEIQSMW KEILLNYGIF LLELLTVFGA IALIVLAIVQ
51  SKKQSESGSV VLTDSENYK KQRQSFETFF LSEEETKHQE KKEKKKEKAE
101 AKAEKKRLKE GGEKSAETQK SRLFLVDFDG DLYAHAVESL RHEITAVLLI
151 AKPEDEVLLR LESPGGVVHG YGLAASQLRR LRERNIPLTV AVDKVAASGG
201 YMMACVADKI VSAPFAVIGS VGVVAEVPNI HRLLKKHDID VDMVTAGEFK
251 RTVTFMGENT EKGKQKFRQE LEETHQLFKQ FVSENRPGLD IEKIATGEHW
301 FGRQALALNL IDEISTSDDL LLKAFENKQV IEVKYQEKRS LIQRIGLQAE
351 ASVEKLFACL VNRRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 21>:

```

m005.seq
1  ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51  GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTCTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AWAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CACAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA

```

```

301 GCAGAGAAAA A.CGTTTGAA GGAGGGTGGG GAGAAATCTG CCGAAACGCA
351 nAAATCACGC CTTTTGTGT TGGANNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGCGAGCGG CGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGCTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTGGTGGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAACAGA AATTCCGACA
801 GGAAGTGGAG GAAACGCATC AGTTGTTCAG GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGCGAA CGGCGGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTTATCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTGCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGT A
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

m005.pep

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQXSGSVVL TDFSENYKKQ RQSFEAFFLS GEEAQHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSAETXKSR LFVLXXXXXX XXXXXXXXXX XXXXXXXXXX
151 XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXASGGYM
201 MACVADKIAS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
301 RQALALNLID EISTDDLLL KAFENKQVIE VKYQEKQSLI QRIQLQAEAS
351 VEKLEAKLVN RRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 23>:

a005.seq

```

1 ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTT CGGAAAATTA
201 TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGAAGAGG
251 CAAAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA
301 GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
351 AAAATCCCGC CTTTTGTGT TGGATTTTGA CGGCGATTG TATGCACACG
401 CCGTAGAATC CTTGCGTCAT GAGATTACGG CCGTGCTTTT GATTGCCAAG
451 CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG GCGTGCTTCA
501 CGGTTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGCGC GAACGCAATA
551 TTCCGCTGAC CGTCGCCGTC GATAAGGTGG CGGCGAGCGG TGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAACAGA AATTCCGACA
801 GGAAGTGGAG GAAACGCATC AGTTGTTCAG GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGCGAA CGGCGGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTTATCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTGCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGTA
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

a005.pep

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQSESGSVVL TDFSENYKKQ RQSFEAFFLS GEEAKHQEKE EKKKEKAEAK
101 AEKKRLKEGG EKSETQKSR LFVLDFDGDY YAHAVESLRH EITAVLLIAK
151 PEDEVLLRLE SPGGVVHGYG LAASQLRRLR ERNIPLTVAV DKVAASGGYM
201 MACVADKIVS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG

```

301 RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIGLQAEAS
 351 VEKLF AKLVN RRADVM*

m005/a005 79.2% identity over a 366 aa overlap

m005.pep	MDNIDMFMPQE	EEIQSMWKE	ILLNYGIFL	LELLTVFGA	I	ALIVLAIVQ	SKKQXSXSGSVL
a005	MDNIDMFMPQE	EEIQSMWKE	ILLNYGIFL	LELLTVFGA	I	ALIVLAIVQ	SKKQSESGSVL
m005.pep	TDFS	ENYKKQRQS	FEAFFLSG	EAAHQH	QEKKEKKKE	KA	EAKA
a005	TDFS	ENYKKQRQS	FEAFFLSG	EAAHQH	QEKKEKKKE	KA	EAKA
m005.pep	LFVL	XX					
a005	LFVL	DFDGLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPGGVVHGYGLAASQLRRLR					
m005.pep	XXXXXXXXXXXXXXXXX	ASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRL	LLKKHDIDVD				
a005	ERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAIVGSVGVVAEVPNIHRL	LLKKHDIDVD					
m005.pep	VMTAGEFKRTVTFMGENT	TEKGKQKFRQ	LEETHQLFKQFVSENRPQLDIEEVATGEHWFG				
a005	VMTAGEFKRTVTFMGENT	TEKGKQKFRQ	LEETHQLFKQFVSENRPQLDIEEVATGEHWFG				
m005.pep	RQALALNLIDEISTSDDLLL	KAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLF	AKLVN				
a005	RQALALNLIDEISTSDDLLL	KAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLF	AKLVN				
m005.pep	RRADVMX						
a005	RRADVMX						

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from *N. gonorrhoeae*:

m005/g005

m005.pep	MDNIDMFMPQE	EEIQSMWKE	ILLNYGIFL	LELLTVFGA	I	ALIVLAIVQ	SKKQXSXSGSV
g005	MGMDNIDMFMPQE	EEIQSMWKE	ILLNYGIFL	LELLTVFGA	I	ALIVLAIVQ	SKKQSESGSV
m005.pep	VL	TDFS	ENYKKQRQS	FEAFFLSG	EAAHQH	QEKKEKKKE	KA
g005	VL	TDFS	ENYKKQRQS	FETFFLSE	EETKHQEKKEKKKE	KA	EAKA

174

	120	130	140	150	160	170
m005.pep	SRLFVLXXX					
			:			
g005	SRLFVLDFDGDLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPGGVVHGYGLAASQLRR					
		130	140	150	160	170
						180
m005.pep	180	190	200	210	220	230
	XXXXXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLKKHDID					
	:					
g005	LRERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAVIGSVGVVAEVPNIHRLKKHDID					
	190	200	210	220	230	240
m005.pep	240	250	260	270	280	290
	VDVMTAGEFKRTVTFMGENTKQKQKFRQLEETHQLFKQFVSENRPQLDIEEVATGEHW					
g005	VDVMTAGEFKRTVTFMGENTKQKQKFRQLEETHQLFKQFVSENRPGLDIEKIATGEHW					
	250	260	270	280	290	300
m005.pep	300	310	320	330	340	350
	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFACL					
g005	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKRSIQRIGLQAEASVEKLFACL					
	310	320	330	340	350	360
m005.pep	360	VNRRADVMX				
g005	VNRRADVMX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 25>:

g006.seq

1	ATGCTGCTGG	TGCTggaatt	ttggttCGGc	gtGtCGGCGG	TGGGCatact
51	tgCGTTGTTT	TTATGGCttt	TGCCACGTTT	TGCCGCCATC	AGCGAAAACC
101	TGTATTTCCG	CCTGAACAAC	AGCTTGGAAC	gcgACAACCA	CTTTATCCGA
151	AAAGGCGACG	AGCGGCAGCT	GTACCGCCAT	TACGGACTGG	TTTCGCGCCT
201	GCGTGTGCTG	ATTTCCAACC	GCGAAGCCTT	CGGCTATCTC	TGCGTCGGCG
251	CGGCGATGGG	TATTTTGTTT	GGCTTTGCTT	TTGTGATGAT	GACGCTCAAA
301	GGCTACGGCA	GCGCGGGGCA	TATTTATTCT	GTCGGCACTT	ATCTGTGGAT
351	GTTTGCCATG	AGTTTGACG	ATGTGCCGCG	ATTGGTCGAA	CAATATCCA
401	ATTTGAAAGA	CATCGGACAA	CGGATAGAGT	GGTCGGAACG	GAACATCAAA
451	GCCGGAACCT	GA			

This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

g006.pep

1	MLLVLEFWFG	VSAVGILALF	LWLLPRFAAI	SENLYFRLNN	SLERDNHFIR
51	KGDERQLYRH	YGLVSRLRVL	<u>ISNREAFGYL</u>	<u>CVGAAMGILF</u>	<u>GFAFVMMTLK</u>
101	GYGSAGHIYS	VGTYLWMFAM	SLDDVPRLVE	QYSNLKDIGQ	RIEWSERNIK
151	AGT*				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 27>:

m006.seq

1	ATGCTGCTGG	TGCTGGAATT	TTGGGTCGGC	GTGTCGGCGG	TGGGCATACT
51	TGCGTTGTTT	TTATGGCTTT	TGCCACGTTT	TGCCGCCATC	AGCGAAAACC
101	TGTATTTCCG	CCTGAACAAC	AGCTTGGAAC	GCGACAACCA	CTTTATCCGA
151	AAAGGCGACC	GGCGGCAGCT	GTACCGCCAT	TACGGACTGC	TTGCGCGCCT
201	GCGTGTGCTG	ATTTCCAACC	GCGAAGCCTT	CGGCTATCTC	TGCGTCGGCA
251	CGGCGATGGG	TATTTTGTTT	GGCTTTGCTT	TTGTGATGAT	GACGCTCAAA
301	GGCTACAGCA	GCGCGGGGCA	TGTCTATTCT	GTCGGCACTT	ATCTGTGGAT

```

351  GTTTGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
451  GCCGGAACCTGA

```

This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>:

```

m006.pep
  1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
 51  KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101  GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGO RIEWSERNIK
151  AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 29>:

```

a006.seq
  1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
 51  TGCCTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101  TGTATTTCCG CCTGAAGAAC AGCTTGAAC GCGACAACCA CTTTATCCGA
151  AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGACTGC TTGCGCGCCT
201  GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
251  CGGCGATGGG TATTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAA
301  GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
351  GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGAAACG GAACATCAAA
451  GCCGGAACCT GA

```

This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

```

a006.pep
  1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERDNHFIR
 51  KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101  GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSNLKDIGO RIEWSKRNIK
151  AGT*

```

m006/a006 96.7% identity over a 153 aa overlap

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS					
a006	MLLVLEFWVGVS					
	10	20	30	40	50	60
m006.pep	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM					
a006	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAI					
	70	80	90	100	110	120
m006.pep	SLDDVPRLVEQYSNLKDIGO RIEWSERNIKAGTX					
a006	SLDDVPRLVEQYSNLKDIGO RIEWSKRNIKAGTX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from *N. gonorrhoeae*:

m006/g006

```

m006.pep      10      20      30      40      50      60
MLLVLEFWVGVS
```

g006-1.seq

1	ATGTGGAAAA	TGTTGAAACA	CATAGCCAAA	ACCCACCGCA	AGCGATTGAT
51	TGGCACATTT	TCCC CGGT CG	GACTGGAAAA	CCTTTTGATG	CTGGGGTATC
101	CGGTGTTTGG	CGGCTGGGCG	ATTAATGCCG	TGATGTCCGG	GAGGTGTGGG
151	CAGGCGTTGC	TGTACGCTTT	GGTTGTATTT	TGTATGTGGC	TGGT CGGTGC
201	GGCACGGCGG	ATTGCCGATA	CGCGCACGTT	TACGCGGATT	TATACCGAAA
251	TCGCCGTGCC	GGTTGTGTTG	GAACAACGGC	AGCGGCAAGT	CCCGCATTCA
301	GCGGTAAC TG	CACGGGTGCG	CCTGTTCGCG	GAATTTTGTA	GCTTTTTTGA
351	AGAACACCTG	CCGATTGCCG	CGACATCCGT	CGTATCCATA	TTCGGCGCGT
401	GCATCATGCT	GCTGGTGCTG	GAATTTTGGG	TCGCGGTGTC	GGCGGTGGGC
451	ATACTTGCGT	TGTTTTTATG	GCTTTTGCCA	CGTTTTGCCG	CCATCAGCGA
501	AAACCTGTAT	TTCCGCTAGT	ACAACAGCTT	GGAACGCGAC	AACCACTTTA
551	TCCGAAAAGG	CGACGAGCGG	CAGCTGTACC	GCCATTACGG	ACTGGTTTTG
601	CGCCTGCGTG	TGCTGATTTC	CAACCGCGAA	CGCTTCGGCT	ATCTCTGCGT
651	CGGCGCGGCG	ATGGGTATTT	TGTTCCGGCT	TGCTTTTGTG	ATGATGACGC
701	TCAAAGGCTA	CGGCAGCGCG	GGGCATATTT	ATTCGGTCGG	CAC TTATCTG
751	TGGATGTTTG	CCATGAGTTT	GGACGATGTG	CCGCGATTGG	TCGAACAATA
801	TTCCAATTTG	AAAGACATCG	GACAACGGAT	AGAGTGGTCG	GAACGGAACA
851	TCAAAGCCGG	AACTTGA			

g006-1.pap

1	1	MWKMLKHIAK	THRKRIGTF	SPVGLENLLM	LGYPVFGGWA	INAVIAGRVM
51	51	QALLYALVVF	LMWLVGARR	IADTRTFTRI	YTEIAPVPVL	EQRQRPVPHS
101	101	AVTARVALSR	EFVSFEEHL	PIAATSVVSI	FGACIMLLVL	EFWVGVSavg
151	151	ILALFLWLLP	RFAAISENLY	FRLNLSLERD	NHFIRKGDGR	QLYRHYGLVS
201	201	RLRLVLSNRE	AFGYLCVGAA	MGILFGFAFV	MMTLKGYGSA	GHIYSVGTYL
251	251	WMFAMSLDDV	PRLVEQYSNL	KDIGQRIEWS	ERNIKAGT*	

m006-1.seq

1	ATGTGGAAAA	TGTTGAAACA	CATAGCCCAA	ACCCACCGCA	AGCGATTGAT
51	TGGCACATTT	TCCCTGGTCG	GACTGGAAAA	CCTTTGTGAT	CTGGTGTATC
101	CGGTGTTTGG	CGCGCGGGCG	ATCAATGCCG	TGATTGCGGG	GGAGGTTGTGG
151	CAGGCGTTTC	TGTACGCTTT	GGTTGTGCTT	TGATGTGGC	TAGTTCGGTG
201	GGTGC GGCG	ATTGCCGATA	CGCGCACGTT	TACGCGGATT	TATACCGAAA
251	TCGCCGTGCC	GGTCGTGTTG	GAACAGCGGC	AGCGACAAGT	CCCGGATTTCG
301	GCGGTAAC TG	CGCGGGTTGC	CCTGTCGCGT	GAGTTTGTCA	GCTTTTTTGA
351	AGAACACCTG	CCGATTGCCG	CGACATCCGT	CGTATCCATA	TTCGGCGCGT
401	GCATCATGCT	GCTGGTGCTG	GAATTTTGGG	TCGGCGTGTC	GGCGCTGGGC
451	ATACTTGCGT	TGTTTTTATG	GCTTTTGCCA	CGTTTTGCCG	CCATCAGCGA
501	AAACCTGTAT	TTCCCGCTGA	ACAACAGCTT	GGAACGCGAC	AACACTTTTA
551	TCCGAAAAGG	CGACCGGCGG	GACGTGTACC	GCCATTACCG	ACTGCTTGCG
601	CGCCTGCGTG	TGCTGATTTC	CAACCGCGAA	GCCTTACGGC	ATCTCTGCGT

```

651  CGGCACGGCG ATGGGTATTT TGTTCCGGCTT TGCTTTTGTG ATGATGACGC
701  TCAAAGGCTA CAGCAGCGCG GGCATGTCT ATTCGGTCGG CACTTATCTG
751  TGGATGTTTG CCATGAGTTT GGACGACGTG CCGCGATTGG TCGAACAATA
801  TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851  TCAAAGCCGG AACTTGA

```

This corresponds to the amino acid sequence <SEQ ID 34; ORF 006-1>:

```

m006-1.pep
  1  MWKMLKHIAQ THRKLIGTF SLVGLNLLM LVYPVFGGRA INAVIAGEVW
 51  QALLYALVVL LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRVPHS
101  AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSavg
151  ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGD RR QLYRHYGLLA
201  RLRVLISNRE AFGYLCVGTA MGILFGFAFV MMTLKGYSSA GHVYSVGTYL
251  WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*

```

m006-1/g006-1 95.5% identity in 288 aa overlap

m006-1.pep	10	20	30	40	50	60
	MWKMLKHIAQTHRKLIGTFSLVGLNLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL					
g006-1	MWKMLKHIAKTHRKLIGTFSPVGLNLLMLGYPVFGGWAINAVIAGRNVQALLYALVVF					
	10	20	30	40	50	60
m006-1.pep	70	80	90	100	110	120
	LMWLVGAVRRIADTRTFTRIYTEIAVPVLEQRQRVPHSAVTARVALSREFVSFFEEHL					
g006-1	LMWLVGAAARRIADTRTFTRIYTEIAVPVLEQRQRVPHSAVTARVALSREFVSFFEEHL					
	70	80	90	100	110	120
m006-1.pep	130	140	150	160	170	180
	PIAATSVVSIFGACIMLLVLEFWVGVSavgILALFLWLLPRFAAISENLYFRLNNSLERD					
g006-1	PIAATSVVSIFGACIMLLVLEFWVGVSavgILALFLWLLPRFAAISENLYFRLNNSLERD					
	130	140	150	160	170	180
m006-1.pep	190	200	210	220	230	240
	NHFIRKGD RR QLYRHYGLLARLRVLISNREAFGYLCVG TAMGILFGFAFVMMTLKGYSSA					
g006-1	NHFIRKGD RR QLYRHYGLVSRRLVLISNREAFGYLCVG AMGILFGFAFVMMTLKGYGSA					
	190	200	210	220	230	240
m006-1.pep	250	260	270	280	289	
	GHVYSVGTYLWMFAMSLDDVPRPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
g006-1	GHIYSVGTYLWMFAMSLDDVPRPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
	250	260	270	280		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 35>:

```

a006-1.seq (partial)
  1  ..AGCCAAAACC ACCGCAAGCG ATTGATTGGC ACATTTTTC TGGTCGGACT
 51  GGAAAACCTT TTGATGCTGG TGTATCCGGT GTTTGGCGGC TGGGCGATTA
101  ATGCCGTGAT TCGGGGCGAG GCGTGGCAGG CGTTGCTGTA CGCTTTGGTT
151  GTGCTTTTGA TGTGGCTGGT CCGTGGCGCG CGGCGGATTG CCGATACGCG
201  CACGTTTACG CGGATTATA CCGAAATCGC CGTGCCGGTT GTGTTGGAAC
251  AGCGGCAGCG GCAAGTCCCG CATTCGGCGG TAACTGCGCG GGTTGCCCTG
301  TCGCGTGAGT TTGTCAGCTT TTTTGAAGAA CACCTGCCGA TTGCCGCGAC
351  ATCCGTGTA TCCATATCG GCGCGTGCAT CATGCTGCTG GTGCTGGAAT
401  TTTGGGTCGG CGTGTGGCGG GTGGGCATAC TTGCGTTGTT TTTATGGCTT
451  TTGCCACGTT TTGCCGCCAT CAGCGAAAAC CTGTATTTC GCCTGAAGAA
501  CAGCTTGGA CGCGACAACC ACTTTATCCG AAAAGGCGAC GAGCGGCAGC
551  TGGACCGCCA TTACGACTG CTTGCGCGCC TGCGTGTGCT GATTTCACAC
601  CGCGAAGCCT TCGGCTATCT CTGCGTCGGC ACGGCGATGG GTATTTTGT
651  CGGCTTTGCT TTTGTGATGA TGACGCTCAA AGGCTACAGC AGCGCGGGGC

```


701 ATGTCTATTC GGTCGGCACT TATCTGTGGA TGTTGCCAT AAGTTTGGAC
 751 GACGTGCCGC GATTGGTCGA ACAATATTCC AATTTGAAAG ACATCGGACA
 801 ACGGATAGAG TGGTCGAAAC GGAACATCAA AGCCGGAAC TGA

This corresponds to the amino acid sequence <SEQ ID 36; ORF 006-1.a>:

a006-1.pep (partial)
 1 ..SQNHRKRLIG TFFLVGLENL LMLVYPVFGG WAINAVIAGQ AWQALLYALV
 51 VLLMWLVGAA RRIADTRTFT RIYTEIAVPV VLEQRQRQVP HSAVTARVAL
 101 SREFVSFFEE HLPAAATSVV SIFGACIMLL VLEFWVGVS VLGILALFLWL
 151 LPRFAAISEN LYFRLKNSLE RDNHFIRKGD ERQLDRHYGL LARLRVLISN
 201 REAFGYLCVG TAMGILFGFA FVMMTLKGYS SAGHVYSVGT YLWMFAISLD
 251 DVPRLVEQYS NLKDIGQRIE WSKRNIKAGT *

a006-1/m006-1 95.7% identity in 280 aa overlap

	10	20	30	40	50
a006-1.pep	SQNHRKRLIGTFFLVGLENLLMLVYPVFGGWAINAVIAGQAWQALLYALVVL				
m006-1	MWKMLKHIAQTHRKLIGTFFSLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL				
	10	20	30	40	50
	60	70	80	90	100
a006-1.pep	LMWLVGAARRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
m006-1	LMWLVGAVRRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
	70	80	90	100	110
	120	130	140	150	160
a006-1.pep	PIAATSVVSIFGACIMLLVLEFWVGVSAGVILALFLWLLPRFAAISENLYFRLKNSLERD				
m006-1	PIAATSVVSIFGACIMLLVLEFWVGVSAGVILALFLWLLPRFAAISENLYFRLNNSLERD				
	130	140	150	160	170
	180	190	200	210	220
a006-1.pep	NHFIRKGDERRQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYS				
m006-1	NHFIRKGDERRQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYS				
	190	200	210	220	230
	240	250	260	270	280
a006-1.pep	GHVYSVGTYLWMFAISLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX				
m006-1	GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX				
	250	260	270	280	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 37>:

g007.seq
 1 atgaACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGcgC
 51 CGCcGCTTCT GCCGccgaca acAGCatcat gaCaAAAGGG CAAAAAGTGT
 101 ACGAATCcAa ctGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
 151 ACTGCGtTTC CTccgctTTT CCggtcgGac tgtattatga acaAACCGCa
 201 cgTCCTgctg cacagcatgg tcaaaggcAt cgacgggaca ttcaaagtgg
 251 agcggcaaaa cctacgacgg atttatgCcc gcaaccgcca tcagcgATGC
 301 GGACATTGCC GCGTCGCCA CTTATATCAT GAACGCCTTT GA

This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>:

g007.pep
 1 MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG
 51 TAFPPLFRSD CIMNKPVLH HSMVKGIDGT FKVERQNLRR IYARNRHQRC
 101 GHCRRLHYH ERL*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 39>:

This corresponds to the amino acid sequence <SEQ ID 40; ORF 007>:

m007.pcp

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 41>:

a007.seq

This corresponds to the amino acid sequence <SEQ ID 42; ORF 007.a>:

a007.pap

m007/a007 97.3% identity over a 113 aa overlap

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 007 shows 86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng) from *N. gonorrhoeae*:

m007/g007

```

      10      20      30      40      50      60
m007.pep  MNTRLPTALVLGCFCAAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD
          |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
g007      MNTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKGEGRGTAFPPFRSD
          10      20      30      40      50      60

      70      80      90     100     110
m007.pep  FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX
          |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

g007 CIMNKPHVLLHSMVKIDGTFKVERQNLRRRIYARNRHQRCGHCRRRHLYHERL
 70 80 90 100 110

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 43>:

g007-1.seq (partial)

```

1  ATGAACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
151 ACTGCGTTTC CTCCGCTTTT CCGGTCGGAC TATATTATGA ACAAACCGCA
201 CGTCCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAGGC AAAAAAAC.
```

This corresponds to the amino acid sequence <SEQ ID 44; ORF 007-1.ng>:

g007-1.pep (partial)

```

1  MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG
51  TAFPPLFRSD YIMNKPHVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKG KKN...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 45>:

m007-1.seq

```

1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAAGC AAAAAAACT
401 AA
```

This corresponds to the amino acid sequence <SEQ ID 46; ORF 007-1>

m007-1.pep

```

1  MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51  TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKS KKN*
```

m007-1 / g007-1 91.7% identity in 133 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCFCAAAS	ADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
g007-1	MNTTRLPTAFILCCLCAAAS	ADNSIMTKGQKVYESNCIACHGKKGEGRG	TAFPPLFRSD			
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPQVLLHSMVKGINGT	IKVNGKTYNGFMPATAISD	ADIAAVATYIMNAFDNGGGSV			
g007-1	YIMNKPHVLLHSMVKGINGT	IKVNGKTYNGFMPATAISD	ADIAAVATYIMNAFDNGGGSV			
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
g007-1	TEKDVQAKGKKN					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 47>:

a007-1.seq (partial)

```

1  ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
```

181

```

151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CCACTGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAAAACA GGCAAAAAAC AAAAAA..

```

This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>:

a007-1.pep (partial)

```

1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKN KK..

```

m007-1/a007-1 98.5% identity in 132 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD				
a007-1	MNTTRLPTALVLGCLCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD				
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV					
a007-1	FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV					
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
a007-1	TEKDVQAKNKK					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 49>:

g008.seq

```

1 ATGAACAACA GACATTTTGC CGTCatcgCC TTGGGCAGCA ACCTTGACAA
51 CCCCGCACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
101 acatccggct tgaaCaggtt tcctcactgt aTatgaccgc acctgtcggg
151 tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCCACCAC
201 CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
251 TCGGACGCGA aCGCAGTTTC CGCAATGCAC CGCGCACATT GGATTTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGGCC TTACCCTGCC
351 GCATCCGCGC GCGCAGGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAATACG GAAAGGTTGT CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GTTAA

```

This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>:

g008.pep

```

1 MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG
51 YDNQPDFINA VCTVSTLDG IALLAELNRI EADFGRRERSF RNAPRTLDD
101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKYGKVELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 51>:

m008.seq

```

1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
51 CCCTGCTCAA CAGGTACGCG CCGCATTTGA CACGCTGTCTG TCCCATCTGT
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT TGTCAATGCC GTCTGCACCG TTTCCACCAC
201 TCTGGACGGC ATTGCCyTGC TTGCCGAAC CAACCGTATC GAGGCTGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GkATTTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACACsCGAC TcACctTGCC

```

```

351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTGTTTTA GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGyTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAATT

```

This corresponds to the amino acid sequence <SEQ ID 52; ORF 008>:

```

m008.pep
  1 MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
 51 YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRRERSF RNAPRTLXLD
101 IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFVL GKHKVAELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 53>:

```

a008.seq
  1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
 51 CCCTGCCCCA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT CGTCAATGCC GTCTGCACCG TTTCCACCAC
201 CTTGGACGGC ATTGCCCTGC TTGCCGAAC CAACCGTATC GAAGCCGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GGATTTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGGAC TCACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGATA AGTAA

```

This corresponds to the amino acid sequence <SEQ ID 54; ORF 008.a>:

```

a008.pep
  1 MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
 51 YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRRERSF RNAPRTLXLD
101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKHKVAELS
151 KRLGNQGIRL LPDK*

```

m008/a008 97.6% identity over a 164 aa overlap

	10	20	30	40	50	60
m008.pep	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYNQPDFVNA					
a008	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYNQPDFVNA					
	10	20	30	40	50	60
m008.pep	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
a008	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDPRLTLPHPR					
	70	80	90	100	110	120
m008.pep	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
a008	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDPRLTLPHPR					
	70	80	90	100	110	120
m008.pep	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
a008	AHERSFVIRPLAEILPDFILGKHGKVAELSKRLGNQGIRLLPDKX					
	130	140	150	160		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from *N. gonorrhoeae*:

```

m008/g008
      10      20      30      40      50      60
m008.pep MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYNQPDFVNA
g008     MNNRHFAVIALGSNLDNPAQQIRGALDALSSHPDIRLEQVSSLYMTAPVGYNQPDFINA

```

183

	10	20	30	40	50	60
	70	80	90	100	110	120
m008.pep	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
g008	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDPRLTLPHPR					
	70	80	90	100	110	120
	130	140	150	160		
m008.pep	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
g008	AHERSFVIRPLAEILPDFILGKYGKVVELSKRLGNQGIRLLPDRX					
	130	140	150	160		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 55>:

g009.seq

```

1  ATGCCCCGCG CTGCCGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAT TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTcg cGTTCCAAGC
201 TGTGTTCAG  GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AaaaGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>:

g009.pep

```

1  MPRAAVAFAER HHHKSKEQN THRRADAEIA EGFAVGNQHT QARNQSVMAV
51  QLPLVAFSDK  VVVAFAQVVQ AEIQVFADGG KTWQKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 57>:

m009.seq

```

1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAT TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTcg CGTTCCAAGC
201 TGTGTTCAG  GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>:

m009.pep

```

1  MPRAAVAFAER HHHKSKEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
51  QLPPVAFSDK  VVVAFAQVVQ AEIQVFADGG KTWQKP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from *N. gonorrhoeae*:

m009/g009

	10	20	30	40	50	60
m009.pep	MPRAAVAFAERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK					
g009	MPRAAVAFAERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARNQSVMAVQLPLVAFSDK					
	10	20	30	40	50	60
	70	80				
m009.pep	VVVAFAQVVQAEIQVFADGGKTWQKPX					
g009	VVVAFAQVVQAEIQVFADGGKTWQKPX					
	70	80				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 59>:

```
a009.seq
  1 ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
 51 CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
201 TGTTCTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:

```
a009.pep
  1 MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
 51 QLPLVAFSDK VVVAFAQVLQ AEIQVFADGG KTWQKP*
```

m009/a009 97.7% identity over a 86 aa overlap

```

              10      20      30      40      50      60
m009.pep      MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK
              |||||
a009           MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK
              10      20      30      40      50      60

              70      80
m009.pep      VVVAFAQVQAEIQVFADGGKTWQKPX
              |||||:|||||
a009           VVVAFAQVLQAEIQVFADGGKTWQKPX
              70      80
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 61>:

```
g010.seq
  1 ATGGGTTTTTCTGTTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
 51 TGCAGGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTGCG GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTGCT TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCGGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGCGCGGAG GGCGGTATTC TGTGAacgc
801 cgacggcgaa cgcTTATGG AAcgctatgc GCcgACCGta aAgaCTTGG
851 CTTCCTCGCga cgtGGTTTCA CgcgcGatgG CGatggaAat ctatgaaggt
901 cgcggctgTG GtaaaAAcaA agaCCacgtC TTAAGTAAAA TCGACcAtAt
951 cggTGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
1001 TTCagtttgc cGGTATCGAT CCGATTAAAG ACCCGATTcc ggttgTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTcGg aCCAATTATC ACGGTGAAGT
1101 TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>:

```
g010.pep
```

```

1  MGFPVRKFDA  VIVGGGGAGL  RAALQLSKSG  LNCAVLISKVF  PTRSHTVAAQ
51  GGISASLGNV  QEDRWDWHMY  DTVKGSDDLW  DQDAIEFMCR  AAPEAVIELE
101 HMGMPFDRVE  SGKIYQRPFG  GHTAEHGKRA  VERACAVADR  TGHAMLHTLY
151 QQNVVRANTQF  FVEWTAQDLI  RDENGDVVGV  TAMEMETGEV  YIFHAKAVMF
201 ATGGGGRIYA  SSTNAYMNTG  DGLGICARAG  IPLEDMFWQ  FHPTGVAGAG
251 VLITEGVRGE  GGILLNADGE  RFMERYAPTV  KDLASRDVVS  RAMAMEIYEG
301 RCGGKNDHV  LLKIDHIGAE  KIMEKLPGIR  EISIQFAGID  PIKDPVPVP
351 TTHYMMGGIP  TNYHGEVVVP  QGDEYEVVK  GLYAAGECAC  ASVHGANRLG
401 TNSLLDLVVF  RPTPR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 63>:

```

m010.seq (PARTIAL)
1  ..nTCCAATTAT  CCAAATCCGG  TCTGAATTGT  GCCGTTTTGT  CTAAAGTGTT
51  CCCGACCCGT  TCGCATACCG  TAGCGGCGCA  GGGCGGTATT  TCCGCCTCTn
101 TGGGTAATGT  GCAGGAAGAC  CGTTGGGACT  GGCACATGTA  CGATACCGTG
151 AAAGGTTCCG  ACTGGTTGGG  CGACCAAGAT  GCGATTGAGT  TTATGTGCCG
201 CGCCGCGCCT  GAAGCCGTAA  TTGAGTTGGA  ACACATGGGT  ATGCCTTTTG
251 ACCGTGTGGA  AAGCGGTAAA  ATTTATCAGC  GTCCTTTCGG  CGGCCATACT
301 GCCGAACACG  GTAAACGCGC  GGTAGAACGC  GYCTGTGCGG  TTGCCGACCG
351 TACAGGTCAT  GCGATGCTGC  ATACTTTGTA  CCAACAAAAC  GTCCGTGCCA
401 ATACGCAATT  CTTTGTGGAA  TGGACGGCAC  AAGATTTGAT  TCGTGATGAA
451 AACGGCGATG  TCGTCGGCGT  AACCGCCATG  GAAATGGAAA  CCGGCGAAGT
501 TTATATTTTC  CACGCTAAAG  CTGTGATGTT  TGCTACCGGC  GCGGCGGGTC
551 GTATTTATGC  GTCTTCTACC  AATGCCTATA  TGAATACCGG  CGATGGTTTG
601 GGTATTTGTG  CGCGTGCAGG  TATCCCGTTG  GAAGACATGG  AATTCTGGCA
651 ATTCCAGCCG  ACCGGCGTGG  CGGGTGCGGG  CGTGTGATT  ACCGAA....

```

This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>:

```

m010.pep (PARTIAL)
1  ..XQLSKSGLNC  AVLSKVFETR  SHTVAAQGGI  SASXGNVQED  RWDWHMYDTV
51  KGSDDLGDQD  AIEFMCRAP  EAVIELEHMG  MPFDRVESGK  IYQRPFGGHT
101 AEHGKRAVER  XCAVADRTGH  AMLHTLYQQN  VRANTQFFVE  WTAQDLIRDE
151 NGDVVGVTAM  EMETGEVYIF  HAKAVMFATG  GGGRIYASST  NAYMNTGDGL
201 GICARAGIPL  EDMFWQFQP  TGVAGAGVLI  TE...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 65>:

```

a010.seq
1  ATGGGCTTTC  CTGTTCCGAA  GTTTGATGCC  GTGATTGTCG  GCGGTGGTGG
51  TGCAGGTTTA  CGCGCANCCC  TCCAATTATC  CAAATCCGGT  CTGAATTGTG
101 CCGTTTGTGC  TAAAGTGTTT  CCGACCCGTT  CGCATACCGT  AGCGGCGCAG
151 GGCGGTATTT  CCGCCTCTCT  GGGTAATGTG  CAGGAAGACC  GTTGGGACTG
201 GCACATGTAC  GATACCGTGA  AAGGTTCCGA  CTGGTTGGGC  GACCAAGATG
251 CGATTGAGTT  TATGTGCCGC  GCCGCGCCTG  AAGCCGTAAT  TGAGTTGGAA
301 CACATGGGTA  TGCCTTTTGA  CCGTGTGGAA  AGCGGTAAAA  TTTATCAGCG
351 TCCTTTCGGC  GGCCATACTG  CCGAACACGG  TAAACGCGCG  GTAGAACGCG
401 CCTGTGCNGT  TGCCGACCGT  ACAGGTCATG  CGATGCTGCA  TACTTTGTAC
451 CAACAAAATG  TCCGTGCCAA  TACGCAATTC  TTTGTGGAAT  GGACGGCACA
501 AGATTTGATT  CGTGATGAAA  ACGGCGATGT  CGTCGGCGTA  ACCGCCATGG
551 AAATGGAAAC  CGGCGAAGTT  TATATTTTCC  ACGCTAAAGC  TGTGATGTTT
601 GCTACCGGCG  GCGGCGGCCG  TATTTATGCG  TCTTCTACCA  ATGCCTATAT
651 GAATACCGGC  GATGGTTTGG  GTATTTGTGC  GCGTGCAGGT  ATCCCGTTGG
701 AAGACATGGA  ATTCTGGCAA  TTCCACCCGA  CCGGCGTGGC  AGGTGCGGGC
751 GTGTTGATTA  CCGAAGGCGT  ACGCGGCGAG  GGCGGTATTC  TGTTGAATGC
801 CGACGGCGAA  CGCTTTATGG  AACGCTATGC  GCCGACCGTA  AAAGACTTGG
851 CTTCTCGCGA  CGTTGTTTCC  CGCGCGATGG  CGATGGAAAT  CTACGAAGGT
901 CGCGGCTGCG  GTAAAAACAA  AGACCATGTC  TTAAGTAAAA  TCGACCATAT
951 CGGCGCAGAA  AAAATTATGG  AAAAAGTACC  GGGCATCCGC  GAGATTTCCA
1001 TTCAGTTCGC  CGGTATCGAT  CCGATTAAAG  ACCCGATTCC  CGTTGTGCCG
1051 ACTACCACT  ATATGATGGG  CGGTATTCCG  ACCAACTACC  ATGGCGAAGT
1101 TGTCGTTTCT  CAAGGCGACG  AATACGAAGT  GCCTGTAAAA  GGTCTGTATG
1151 CGGCAGGTGA  GTGCGCCTGT  GCTTCCGTAC  ACGGTGCGAA  CCGCTTGGGT
1201 ACGAACTCCC  TGCTGGACTT  AGTGGTATTC  GGTAAAGCTG  CCGGCGACAG
1251 CATGATTAAG  TTCATCAAAG  AGCAAAGCGA  CTGAAACCT  TTGCCTGCTA

```


186

```

1301 ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGA CAATCAAAC
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GCGGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCACGCGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

```

a010.pep
  1 MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKV FPTRSHTVAAQ
 51 GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLSR DVVS RAMAMEIYEG
301 RCGGNKNDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMAIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010/a010 98.7% identity over a 231 aa overlap

```

                                10      20      30
m010.pep                      XQLSKSGLNCVLSKVFPTRSHTVAAQGGISASXGNV
                                |||
a010      MGFPVRKFDAVIVGGGGAGLRAXLQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m010.pep      QEDRWDWHMYDTVKGSDWLG DQDAIEFMCR AAPEAVIELEHMGMPFDRVESGKIYQRPFG
                                |||
a010      QEDRWDWHMYDTVKGSDWLG DQDAIEFMCR AAPEAVIELEHMGMPFDRVESGKIYQRPFG
                                70      80      90      100     110     120

                                100     110     120     130     140     150
m010.pep      GHAEHGKRAVERXC AVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V
                                |||
a010      GHAEHGKRAVERAC AVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V
                                130     140     150     160     170     180

                                160     170     180     190     200     210
m010.pep      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                                |||
a010      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                                190     200     210     220     230     240

                                220     230
m010.pep      FQPTGVAGAGVLITE
                                |:|
a010      FHPTGVAGAGVLITEGVREGGILLNADGERFMERYAPT VKDLSR DVVS RAMAMEIYEG
                                250     260     270     280     290     300

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from *N. gonorrhoeae*:

m010.pep/g010.pep

```

                                10      20      30
m010.pep                      XQLSKSGLNCAVL SKVFPTRSH TVAAQGGISASXGNV
                                |||||
g010      MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVL SKVFPTRSH TVAAQGGISASLGNV
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m010.pep      QEDRW DWHMYD TVKGS DWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                                |||||
g010      QEDRW DWHMYD TVKGS DWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
                                70      80      90      100     110     120

                                100     110     120     130     140     150
m010.pep      GH TAEHGKRAVERXC AVADRTGHAM LHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVGV
                                |||||
g010      GH TAEHGKRAVERAC AVADRTGHAM LHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVGV
                                130     140     150     160     170     180

                                160     170     180     190     200     210
m010.pep      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                                |||||
g010      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
                                190     200     210     220     230     240

                                220     230
m010.pep      FQPTGVAGAGVLITE
                                |:|||||
g010      FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPT VKDLASRDVVS RAMAMEIYEG
                                250     260     270     280     290     300

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 67>:

g010-1.seq..

```

1   ATGGGTTTTTCTGTTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
51  TGCAGGTTTACGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTTTGTC TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GCGGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTGCT TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTGCGC CCGTGGCGGC ATTCGGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGGCGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GGCGGTATTC TGTGAACGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTGGTTTCA CGCGCGATGG CGATGGAAAT CTATGAAGGT
901 CGCGGCTGTG GTAAAAACAA AGACCACGTC TTACTGAAAA TCGACCATAT
951 CCGTGACAGAA AAAATTATGG AAAAATGCC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTTGC CGGTATCGAT CCGATTAAAG ACCCGATTCC GGTGTGCCC
1051 ACTACCACT ATATGATGGG CGGCATTCCG ACCAATTATC ACGGTGAAGT
1101 TGTGTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT

```

1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTT cgcccaaccc cccggtga

This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:

g010-1.pep

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVFPTRSHTVAAQ
51  GGISASLGNV QEDRWDDHMY DTVKGSDDLW DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHATAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRRANTQF FVEWTAQDLI RDENGDDVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLASRDVVS RAMAMEIYEG
301 RGCCKNKDHF LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDPVPVP
351 TTHYMMGGIP TNYHGEVVPV QGDEYEVVPV GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

g010-1 / P10444

sp|P10444|DHSA_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT
 gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein [Escherichia coli] gi|1786942
 (AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588
 Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 191/303 (63%), Positives = 238/303 (78%)

```

Query: 1  MGFPVRKFDAVIVXXXXXXXXXXXXXSKSLNCVLSKVFPTRSHTVAAQGGISASLGNV 60
      M  PVR+FDAV++ S+SG CA+LSKVFPTRSHTV+AQGGI+ +LGN
Sbjct: 1  MKLPVREFDAVVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60

Query: 61 QEDRWDDHMYDTVKGSDWLGDDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG 120
      ED W+WHMYDTVKGSD++GDQDAIE+MC+ PEA++ELEHMG+PF R++ G+IYQRPFG
Sbjct: 61 HEDNWEWHMYDTVKGSDYIGDQDAIEYMCKTGPEAILELEHMGLPFSRLDDGRIYQRPFG 120

Query: 121 GHATAEHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDDVGV 180
      G + G R A ADRTGHA+LHTLYQQN++ +T F EW A DL++++G VVG
Sbjct: 121 GQSKNFGGEQAARTAAADRTGHALLHTLYQQNLKVNHTTIFSEWYALDLVKNQDQAVVGC 180

Query: 181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ 240
      TA+ +ETGEV F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ
Sbjct: 181 TALCIETGEVVYFKARATVLATGGAGRIYQSTTNAHINTGDGVGMAIRAGVPVQDMEMWQ 240

Query: 241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPT VKDLASRDVVS RAMAMEIYEG 300
      FHPTG+AGAGVL+TEG RGEAG LLN GERFMERYAP KDLA RDVV+R++ +EI EG
Sbjct: 241 FHPTGIAGAGVLTEGCRGEGGYLLNKHGERFMERYAPNAKDLAGRDVVARSIMIEIREG 300

Query: 301 RGC 303
      RGC
Sbjct: 301 RGC 303

```

Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 53/102 (51%), Positives = 62/102 (60%)

```

Query: 309 HVLLKIDHIGAEKIMEKLPGIREISIQFAGXXXXXXXXXXXXTTHYMMGGIPTNYHGEVV 368
      H LK+DH+G E + +LPGI E+S FA T HYMMGGIPT G+ +
Sbjct: 310 HAKLKLHDLGKEVLESRLPGILELSRTFAHVDPVKEPIPIPTCHYMMGGIPTKVTGQAL 369

Query: 369 VPQGDEYEVVPVGLYAAGECACASVHGANRLGTNSLLDLVVF 410
      +V V GL+A GE AC SVHGANRLG NSLLDLVVF
Sbjct: 370 TVNEKGEDVVVPGLFVAGEIACVSVHGANRLGGNSLLDLVVF 411

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 69>:

m010-1.seq..

```

1  ATGGGTTTTTCTGTTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTACGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTCTAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GCGCGGTATTTCCGCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTACGATACCGTGA AAGGTTCCGA CTGTTTGGGC GACCAAGATG
251 CGATTGAGTTTATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC

```

```

451 CAACAAAACG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGTCTG TATTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGCGTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTAGTGAAAA TCGACCATAT
951 CCGGCGCAGAA AAAATTATGG AAAAAGTACC GGGCATCCGC GAGATTTCCT
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTCCG ACCAATTACC ACGGCGAAGT
1101 TGTCGTTCCG CAAGGTGAAG ATTACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCTTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACCAACTCCC TGTGGACTT GGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301 ATGCAGGTGA GTTGACCCGC CAACGTATCG AGCGTTTGA CAACCAAACC
1351 GATGGTGAAG ACGTTGATGC ATTGCGTCGC GAACTGCAAC GCTCTGTACA
1401 ACTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTGG AATTGGATAA
1551 CCTGATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCAGCGCG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATATCAATA CTTGTCTCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGATGA

```

This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>:

m010-1.pep..

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKV F PTRSHTVAAQ
51  GGISASLGNV QEDRWDDHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGV RGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RCGGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGEDYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMATA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAAH SDDHPPERDDE
551 NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010-1 / g010-1 99.5% identity in 410 aa overlap

	10	20	30	40	50	60
m010-1.pep	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV					
g010-1	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV					
	10	20	30	40	50	60
m010-1.pep	QEDRWDDHMYDTVKGSDWLGDDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG					
g010-1	QEDRWDDHMYDTVKGSDWLGDDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG					
	70	80	90	100	110	120
m010-1.pep	GHAEHGKRAVERACAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVG V					
g010-1	GHAEHGKRAVERACAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVG V					
	130	140	150	160	170	180
m010-1.pep	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
g010-1	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
	190	200	210	220	230	240

190

	190	200	210	220	230	240
	250	260	270	280	290	300
m010-1.pep	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
g010-1	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
	250	260	270	280	290	300
	310	320	330	340	350	360
m010-1.pep	RGCCKNDHVLKIDHIGA EKIMEKLP GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP					
g010-1	RGCCKNDHVLKIDHIGA EKIMEKLP GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP					
	310	320	330	340	350	360
	370	380	390	400	410	420
m010-1.pep	TNYHGEVVVPQGDEYEVVVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK					
g010-1	TNYHGEVVVPQGDEYEVVVKGLYAAGECACASVHGANRLGTNSLLDLVVFRPTPRX					
	370	380	390	400	410	
	430	440	450	460	470	480
m010-1.pep	FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTDEILS					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 71>:

a010-1.seq..

1	ATGGGCTTTC	CTGTTCGCAA	GTTTGATGCC	GTGATTGTGCG	GCGGTGGTGG
51	TGCAGGTTTA	CGCGCANCCC	TCCAATTATC	CAATCCGGT	CTGAATTGTG
101	CCGTTTGTGC	TAAAGTGTTC	CCGACCCGTT	CGCATACCGT	AGCGGCGCAG
151	GGCGGTATTT	CCGCCTCTCT	GGGTAATGTG	CAGGAAGACC	GTGGGACTG
201	GCACATGTAC	GATACCGTGA	AAGGTCCGA	CTGGTTGGGC	GACCAAGATG
251	CGATTGAGTT	TATGTGCCGC	GCCGCGCCTG	AAGCCGTAAT	TGAGTTGGAA
301	CACATGGGTA	TGCCTTTGA	CCGTGTGGAA	AGCGGTAAAA	TTTATCAGCG
351	TCCTTTCGGC	GGCCATACTG	CCGAACACGG	TAAACGCGCG	GTAGAACGCG
401	CCTGTGCNGT	TGCCGACCGT	ACAGGTCATG	CGATGCTGCA	TACTTTGTAC
451	CAACAAATG	TCCGTGCCAA	TACGCAATTC	TTTGTGGAAT	GGACGGCACA
501	AGATTTGATT	CGTGATGAAA	ACGGCGATGT	CGTCGGCGTA	ACCGCCATGG
551	AAATGGAAAC	CGGCGAAGTT	TATATTTTCC	ACGCTAAAGC	TGTGATGTTT
601	GCTACCGGCG	GCGGCGGCCG	TATTTATGCG	TCTTCTACCA	ATGCCTATAT
651	GAATACCGGC	GATGGTTTGG	GTATTTGTGC	GCGTGCAGGT	ATCCCGTTGG
701	AAGACATGGA	ATTCTGGCAA	TTCCACCCGA	CCGCGCTGGC	AGGTGCGGGC
751	GTGTTGATTA	CCGAAGGCGT	ACGCGGCGAG	GGCGGTATTC	TGTTGAATGC
801	CGACGGCGAA	CGCTTTATGG	AACGCTATGC	GCCGACCGTA	AAAGACTTGG
851	CTTCTCGCGA	CGTTGTTTCC	CGCGCGATGG	CGATGGAAAT	CTACGAAGGT
901	CGCGGCTGCG	GTAAAAACAA	AGACCATGTC	TTACTGAAAA	TCGACCATAT
951	CGGCGCAGAA	AAAATTATGG	AAAAACTGCC	GGGCATCCGC	GAGATTTCCA
1001	TTCAGTTCCG	CGGTATCGAT	CCGATTAAAG	ACCCGATTCC	CGTTGTGCCG
1051	ACTACCCACT	ATATGATGGG	CGGTATTCCG	ACCAACTACC	ATGGCGAAGT
1101	TGTCGTTTCT	CAAGGCGACG	AATACGAAGT	GCCTGTAAAA	GGTCTGTATG
1151	CGGCAGGTGA	GTGCGCCTGT	GCTTCCGTAC	ACGGTGCAGG	CCGCTTGGGT
1201	ACGAACCTCC	TGCTGGACTT	AGTGGTATTC	GGTAAAGCTG	CCGCGCACAG
1251	CATGATTAAA	TTCATCAAAG	AGCAAAGCGA	CTGGAAACCT	TTGCCTGCTA
1301	ATGCCGGCGA	ACTGACCCGC	CAACGTATCG	AGCGTTTGGA	CAATCAAACT
1351	GATGGTGAAA	ACGTTGATGC	ATTGCGCCGC	GAACGTCAAC	GCTCCGTACA
1401	ATTGCACGCC	GGCGTGTTC	GTAATGATGA	GATTCTGAGC	AAAGGCGTTC
1451	GAGAAGTCAT	GGCGATTGCC	GAGCGTGTGA	AACGTACCGA	AATCAAAGAC
1501	AAGAGCAAAG	TGTGGAATAC	CGCGCGTATC	GAGGCTTTGG	AATTGGATAA
1551	CGTAATTGAA	GTGGCGAAAAG	CGACTTTGGT	GTCTGCCGAA	GCACGTAAAG
1601	AATACGCGG	TGCGCACGCT	TCAGACGACC	ATCCTGAGCG	CGATGATGAA
1651	AACTGGATGA	AACATACGCT	GTACCATTCA	GATGCCAATA	CCTTGTCCCTA
1701	CAAACCGGTG	CACACCAAGC	CTTTGAGCGT	GGAATACATC	AAACCGGCCA
1751	AGCGCGTTTA	TTGA			

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>:

a010-1.pep..

1	MGFPVRKFDA	VIVGGGGAGL	RAXLQLSKSG	LNCVLSKVF	PTRSHTVAAQ
51	GGISASLGNV	QEDRDWHMY	DTVKGSDWL	GQDAIEFMCR	AAPEAVIELE
101	HMGMPFDRVE	SGKIYQRPFG	GHTAEHGKRA	VERACAVADR	TGHAMLHTLY

151 QQNVRANTQF FVEWTAQDLI RDENGDVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCGKNKDHV LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKDPVPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWK LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMAIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*

m010-1 / a010-1 99.3% identity in 587 aa overlap

	10	20	30	40	50	60
a010-1.pep	MGFPVRKFD	AVIVGGGGAG	LRLXQLSKS	GSLNCAVLS	KVFPTRSH	TVAAQGGISAS
a010-1	MGFPVRKFD	AVIVGGGGAG	LRLAQLSKS	GSLNCAVLS	KVFPTRSH	TVAAQGGISAS
	10	20	30	40	50	60
a010-1.pep	QEDRDW	WHMYDTVK	GS	DWLGDQDA	IEFMCRAA	PEAVIELEHM
m010-1	QEDRDW	WHMYDTVK	GS	DWLGDQDA	IEFMCRAA	PEAVIELEHM
	70	80	90	100	110	120
a010-1.pep	GHTAEHG	KRAVERAC	AVADRTGH	AMLHTLYQ	QNV	RANTQFFVEW
m010-1	GHTAEHG	KRAVERAC	AVADRTGH	AMLHTLYQ	QNV	RANTQFFVEW
	130	140	150	160	170	180
a010-1.pep	TAMEMET	GEVYIFHA	KAVMFATG	GGGGRIYA	SSTNAYMN	TG
m010-1	TAMEMET	GEVYIFHA	KAVMFATG	GGGGRIYA	SSTNAYMN	TG
	190	200	210	220	230	240
a010-1.pep	FHPTGV	AGAGVLIT	EGVRGEGG	ILLNADGE	RFMERYAP	TVKDLASR
m010-1	FHPTGV	AGAGVLIT	EGVRGEGG	ILLNADGE	RFMERYAP	TVKDLASR
	250	260	270	280	290	300
a010-1.pep	RGCGKN	KDHVLLK	IDHIGAEK	IMEKLP	GIREISIQ	FAGIDPIK
m010-1	RGCGKN	KDHVLLK	IDHIGAEK	IMEKLP	GIREISIQ	FAGIDPIK
	310	320	330	340	350	360
a010-1.pep	TNYHGE	VVPVQGE	YEV PVK	GLYAAGE	CACASVH	GANRLGT
m010-1	TNYHGE	VVPVQGE	YEV PVK	GLYAAGE	CACASVH	GANRLGT
	370	380	390	400	410	420
a010-1.pep	FIKEQSD	WKPLPAN	AGELTRQ	RIERLDN	QTDGENV	DALRRELQ
m010-1	FIKEQSD	WKPLPAN	AGELTRQ	RIERLDN	QTDGENV	DALRRELQ
	430	440	450	460	470	480
a010-1.pep	KGVREV	MAIAER	VKRTEIK	DKSKVW	NTARIEA	EALDNLIE
m010-1	KGVREV	MAIAER	VKRTEIK	DKSKVW	NTARIEA	EALDNLIE
	490	500	510	520	530	540
a010-1.pep	SDDHPE	RDDENW	MKHTLY	HS	DANTLSY	KPVHTKPL
m010-1	SDDHPE	RDDENW	MKHTLY	HS	DANTLSY	KPVHTKPL
	550	560	570	580		
a010-1.pep	SDDHPE	RDDENW	MKHTLY	HS	DANTLSY	KPVHTKPL
m010-1	SDDHPE	RDDENW	MKHTLY	HS	DANTLSY	KPVHTKPL

m010-1 SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX
550 560 570 580

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 73>:

g011.seq
1 ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
101 GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA
151 GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAACA
201 GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
251 TCCTGACCAA AATGGTCAAA CAGCGCAAAG ACGGCGCGAA AATCTACACT
301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGACGT
351 GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG
401 CCGTCGAAGC AGCCGTGGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
451 GGCAAAGTGA TGGTCGTATT GAAAACccGC CTCGCCGGCA AAGccgATAT
501 GGGCGAAGTC AACAAAATCT TGAAAACcGt aCTGACCGCC tga

This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:

g011.pep
1 MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK
51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAKIYT
101 EAGRQDLADK ENAEIDVLHR YLPQMLSAGE IRTAVEAAVA ETGAAGMADM
151 GKVMVVLKTR LAGKADMGEV NKILKTVLTA *

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 75>:

m011.seq (partial)
1 ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
101 GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA
151 GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAACA
201 GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
251 TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACAGCGCGAA AATCTACACT
301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGAGGT
351 ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG
401 AGGTCAAGC TGCCGTGGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
451 GGTAAAGTCA TGGGGCTGCT GAAAACCGCG CTCGCAGGTA AAGCCGA...

This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:

m011.pep (partial)
1 MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK
51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKIYT
101 EAGRQDLADK ENAEIEVLHR YLPQMLSAGE IRTEVEAAVA ETGAAGMADM
151 GKVMGLLKTR LAGKA.....

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng) from *N. gonorrhoeae*:

m011/g011

	10	20	30	40	50	60
m011.pep	MRTHRKTC	SAVCFAFQ	TASKPAVS	IRHPSEDI	MSLKIRLT	EDMKTAMRAKDQVSLGTIRL
	:	:	:	:	:	:
g011	MKTHRKTC	SAVCFAFQ	TASKPAVS	IRHPSEDI	MSLKTRLT	EDMKTAMRAKDQVSLGTIRL

	70	80	90	100	110	120
m011.pep	INAAVKQFEV	DERTEADDAK	ITAILTKMVK	QRKDSAKIYT	EAGRQDLADK	ENAEIEVLHR
g011	INAAVKQFEV	DERTEADDAK	ITAILTKMVK	QRKDGAKIYT	EAGRQDLADK	ENAEIDVLHR

193

	70	80	90	100	110	120
	130	140	150	160		
m011.pep	YLPQMLSAGEIRTEVEAAVAETGAAGMADMGKVMGLLKTRLAGKA					
	:					
g011	YLPQMLSAGEIRTAVEAAVAETGAAGMADMGKVMVVLKTRLAGKADMGEVNKILKTVLTA					
	130	140	150	160	170	180
g011	X					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 77>:

g012.seq

```

1  ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTCAC
51  TGACAAACTG CTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCACa
201 gGcggTGGAT ATTCGgcact tccgCcacca caccaccga accgatgacc
251 gcaaacggaG CGGAAACAAT TTTATCCGcC acacacgcca tcatatagcc
301 gcCGCTTGCC GCACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCCTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGCCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATT TCGCTTTCG
551 GCAGATTTCT CCCCACCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CTTTTTTTTC CTGATGTTT GTCTCTTCTC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>:

g012.pep

```

1  MLARRYFFNI QPGAVETDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRHFRHHTHR TDDRKRSGNN FIRHTRHHIA
101 AACRDLIDGD GQRNIAFAQT PKLSRQTVT VNHAARTFQS EQNLIFRLGN
151 QKHRNLMTO GFYGVCIQIA VKIQHKKAGF LRFRGRFLPAL LQTLFLCFGF
201 RLFLFLFFFF LMFCLFLA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 79>:

m012.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCACa
201 GGCGGTGGAT ATTCGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
351 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
451 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
501 nnnnnnnnnn nnnnnnnnnC AACACAAAAA GGCGTGATT nTGCCTTTCG
551 GCAGATTTCT CCCCACCTC CTTCAAACGT TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTGT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>:

m012.pep

```

1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRHHTHR TDNRKRSGSN FIRHTRHHIT
101 AARXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XXXQHKKA*F XRFRGRFLPTL LQTFFLCFGF
201 RLFLFLFLFF LMCLFLPA*

```


The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 81>:

```
a012.seq
1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GCGGTGGAT ATTGCTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCCTAACG GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GAAGATTCTT CCCCACCCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:

```
a012.pep
1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKHRRLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLFPA*
```

m012/a012 64.2% identity over a 218 aa overlap

	10	20	30	40	50	60
m012.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
a012	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m012.pep	NIMFFQQA VDIRYFRHHHTDNRKRSGNFIRHTRHHITAARXXXXXXXXXXXXXXXXXXXX					
a012	NIMFFQQA VDIRYFRYNTHRTDNRKRSGNFIRHTRHHITTARRHLIDGDGQRNIAFAQT					
	70	80	90	100	110	120
	130	140	150	160	170	180
m012.pep	XXXQHKKAXF					
a012	PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRLMTQGFYGVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180
	190	200	210	219		
m012.pep	XRFGRLPTLLQTFFLCFGFRLFLFLFLFLFLMLCLFPAX					
a012	LRFGRLPTLLQTLFLCFGFRLFLFLFLFLFLMLFCLFPAX					
	190	200	210			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng) from *N. gonorrhoeae*:

m012/g012

	10	20	30	40	50	60
m012.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					

195

	10	20	30	40	50	60
m012.pep	70	80	90	100	110	120
	NIMFFQQAVIDIRYFRHHHTHRTDNRKRSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXXXXX					
	: : : : : :					
g012	70	80	90	100	110	120
	NIMFFQQAVIDIRHFRHHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLIDGDGQRNIAFAQT					
	130	140	150	160	170	180
m012.pep	XXQHKKAXF					
	:	:	:	:	:	
g012	130	140	150	160	170	180
	PKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKRRNLMTQGFGVCIQIAVKIQHKKAGF					
	190	200	210	219		
m012.pep	XRFGFRFLPTLLQTFFLCFGRFLFLFLFFLMLCLFPAX					
	: : : : : :					
g012	190	200	210			
	LRFGFRFLPALLOTFLCFGRFLFLFLFFFLMFCFLAX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 83>:

m012-1.seq

```

1   ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CyTAAGCTGC GAAGCCGCCA AACCCTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATT TCGCGTTTCG
551 GCAGATTTCT CCCACCTTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 84; ORF 012-1>:

m012-1.pep

```

1   MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVID IRYFRHHHTH TDNRKRSGSN FIRHTRHHIT
101 AARRHLIDGD GQRNIAFAQT XKLRSRQTVT VNHAARTFQS EQNLIFRLGN
151 QKRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGFRFLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLEFPA*

```

m012-1/g012 91.7% identity in 218 aa overlap

	10	20	30	40	50	60
m012-1.pep	MLARCHFLNIQLRAVLADKLLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
	: : : : : : :					
g012	10	20	30	40	50	60
	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					
	70	80	90	100	110	120
m012-1.pep	NIMFFQQAVIDIRYFRHHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT					
	: : : : : : : : : :					
g012	70	80	90	100	110	120
	NIMFFQQAVIDIRHFRHHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLIDGDGQRNIAFAQT					
	130	140	150	160	170	180
m012-1.pep	XKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKRRNLMTQGFGVCIQIAVKIQHKKAGF					
	: : : : : : : : : :					
g012	PKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKRRNLMTQGFGVCIQIAVKIQHKKAGF					

196

	130	140	150	160	170	180		
	190	200	210	219				
m012-1.pep	LRFG	RFLP	TLQL	FLFC	FGFR	FLFL	FLFF	LMFCLFPAX
g012	LRFG	RFLP	ALLQ	TLFL	CFGR	FLFL	FLFF	LMFCLFLAX
	190	200	210					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 85>:

a012-1.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTGGAACAAC TGATGCGTTT CCTCCAGTTC CTGTGCGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCCTT
151 AAATTGCCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTCGGTACT TCCGCTACAA CACCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTACGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GAAGATTTCCT CCCACCCCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC GTGATGTTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:

a012-1.pep

```

1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSNN FIRHTRHHIT
101 TARRHLIDGD QORNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLFPA*

```

a012-1/m012-1 97.2% identity in 218 aa overlap

	10	20	30	40	50	60		
a012-1.pep	MLARCHFLNIQLRAVLADKLL	LEQLMRFLQFLSEFLFALFR	IFTHKSNRALKFARRHHIHI					
m012-1	MLARCHFLNIQLRAVLADKLL	LEQLMRFLQFLSEFLFALFR	IFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60		
	70	80	90	100	110	120		
a012-1.pep	NIMFFQQAVDIRYFRYNTHRT	DNKRKSGNNFIRHTRHHIT	TARRHLIDGDGQORNIAFAQT					
m012-1	NIMFFQQAVDIRYFRHHTHRT	DNKRKSGSNFIRHTRHHITA	ARRHLIDGDGQORNIAFAQT					
	70	80	90	100	110	120		
	130	140	150	160	170	180		
a012-1.pep	PKLRSRQTVT	VNHAARTFQSKQNLIFRLGN	QKHRRNLMTQGFYGVCIQIA	VKIQHKKAGF				
m012-1	XKLSRQTVT	VNHAARTFQSEQNLIFRLGN	QKHRRNLMTQGFYGVCIQIA	VKIQHKKAGF				
	130	140	150	160	170	180		
	190	200	210	219				
a012-1.pep	LRFG	RFLP	TLQL	FLFC	FGFR	FLFL	FLFF	LMFCLFPAX
m012-1	LRFG	RFLP	TLQL	FLFC	FGFR	FLFL	FLFF	LMFCLFPAX
	190	200	210					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 87>:

g013.seq

```

1  aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTtgt tcataataca
51  gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGCGCCCT TCCCCTTTCT
101 TGCCGTGGCA GGCGATGcag tTgGATTcGT AACTTTTTG CCCTTTtGtc

```

197

```

151 atgatGCTgt tgtcggCGGC AGAAGCgGCG GcgCAGAGGC AGCACAAGAT
201 GAAGGCGGTC GGCAGTCGGG TTGTGTtcat tGgcgTTTCC cctaattgtt
251 tgaaaccttg ttttttgatt Ttgcctttac ggggtgaaaa gtttttTtgg
301 cccaaatccg gaatttag

```

This corresponds to the amino acid sequence <SEQ ID 88; ORF 013.ng>

g013.pep

```

1 MPLTMLCSRT CGLFIIQSDR KSGGNAVPRP SPFLPWQAMQ LDSYTFPCFV
51 MMLLSAAEAA AQRQHMKAV GSRVVFIVGS PNVLKPCFLI LPLRGEKFFW
101 PKSGI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 89>:

m013.seq

```

1 ATGCCTTTGA CCATGCTGTG CAGCAGCACC TCGGTTTTT TCATGATGAA
51 GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101 TGCCGTGGCA GGCACGCAG TTGGATTCGT AACTTTTTG CCCTTTTGTC
151 ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAAGC AGCCCAAGAC
201 GAGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTC TCCATGTTTG
251 AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT TTGCGCAAT
301 CAGTAG

```

This corresponds to the amino acid sequence <SEQ ID 90; ORF 013>:

m013.pep

```

1 MPLTMLCSST CGFFMMKSER XSGGNMVRP SPFLPWQATQ LDSYTFPCFV
51 MMLLSAAEAA AQKQPKTRAV GSRVVFIVGS FMFETLLLLIL RSGXKIFLPN
101 Q*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 91>:

a013.seq

```

1 ATGCCTTTGA CCATGCTGTG CAGCAGCACC TCGGTTTTT TCATGATGAA
51 GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101 TGCCGTGGCA GGCACGCAG TTGGATTCGT AACTTTTTG CCCTTTTGTC
151 ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAGGC AGCCCAAGAC
201 GAGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTC TTAATGTTTG
251 AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT CTTGCCGAAT
301 CGGTAG

```

This corresponds to the amino acid sequence <SEQ ID 92; ORF 013.a>:

a013.pep

```

1 MPLTMLCSST CGFFMMKSER *SGGNMVRP SPFLPWQATQ LDSYTFPCFV
51 MMLLSAAEAA AQRQPKTRAV GSRVVFIVGS LMFETLLLLIL RSG*KIFLPN
101 R*

```

m013/a013 97.0% identity over a 101 aa overlap

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFPCFVMMMLLSAAEAA					
a013	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFPCFVMMMLLSAAEAA					
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVFIVGSFMFETLLLLILRSGXKIFLPNQX					
	:					
a013	AQRQPKTRAVGSRVVFIVGSLMFETLLLLILRSGXKIFLPNRX					
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from *N. gonorrhoeae*:

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 93>:

This corresponds to the amino acid sequence <SEQ ID 94; ORF 015.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 95>:

This corresponds to the amino acid sequence <SEQ ID 96; ORF 015:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 97>:

This corresponds to the amino acid sequence <SEQ ID 98; ORF 015.a>:

a015.pep

199

```

1  MQYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKPLAGFW KALPHLNDTM
51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMM RARPRSTKFY
101 TVYLLAMCCL TCIVYLAKTK VLPF*

```

m015/a015 96.7% identity over a 91 aa overlap

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | | | | | |
a015      LIVKYSHQIFVTITILVFNIRVFXLWKNPEKPLAGFWKALPHLNDTMLLFTGLWLMKITH
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m015.pep      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFTYTVYLLAMCCVACIVYLAKTKVLP
                | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a015      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFTYTVYLLAMCCLTCIVYLAKTKVLP
                                70      80      90      100     110     120

m015.pep      FX
                ||
a015          FX

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from *N. gonorrhoeae*:

m015/g015

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | | | | | |
g015      LIVKYSHQIFVTITILVFNIRFLLWKNPEKPLVGFVKALPHLNDTMLLFTGLWLMKITH
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m015.pep      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFTYTVYLLAMCCVACIVYLAKTKVLP
                | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g015      FSPFNAPWLGTKILLLFAYIALGMVMRRARPRSTKFTYTVYLLAMCCIIACIVYLAKTKVLP
                                70      80      90      100     110     120

m015.pep      FX
                ||
g015          FX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 99>:

g018.seq

```

1  atGCAGCAGG GGCagttggt tggacgcgtc gcccgcaata AAGATATGCG
51  GAATgctggt CTGCATggtC AGCGGATCGG CAACGGGtac gccgcgcgcg
101 tctttgTCGA TATTGATGTT TTCAAACCG ATATtgTCAA CGTTCGGACG
151 GCgACCTACG GCTGCCAACA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT tctACATTGC CGTCTGCGTC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT tctTCTCCGA ACACGGCTTT
301 CGCCTCGTCT GA

```

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>:

g018.pep

```

1  MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVFVDIDV FQTDIVNVRT
51  ATYGQCHIFG NKYAFAILL PMDFYIAVCV EFDLGFSIQM QFQFFSEHG

```

101 RLV*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 101>:

```
m018.seq
  1 ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
 51 GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATATT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>:

```
m018.pep
  1 MQQRQLVGRV ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT
 51 AAHGQHIIFG NKYAFAILL PMDFYIAVCI EFDLGFSIQM QFQFFAEHGV
101 RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 103>:

```
a018.seq
  1 ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
 51 GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCTACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGGCC
251 TCGGTTTTAG CATCCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

```
a018.pep
  1 MQQQQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT
 51 AAYGQHIIFG NKYAFAILL PMDFYIAVCV EFDLGFSIQM QFQFFTEHGF
101 RLV*
```

m018/a018 86.4% identity over a 103 aa overlap

	10	20	30	40	50	60
m018.pep	MQQRQLVGRV	ACDEDMRNTG	LHGQRVGNRY	AARIFFDIDI	FQTDIVNVRT	AAHGQHIIFG
		:	:	:	:	:
a018	MQQQQLVGRV	ARNKDMRNTG	LHSQRIGNGY	AARIFFDIDV	FQTDIVNVRT	AAYGQHIIFG
	10	20	30	40	50	60
	70	80	90	100		
m018.pep	NKYAFAILL	PMDFYIAVCI	EFDLGFSIQM	QFQFFAEHGV	RLVX	
		:	:	:		
a018	NKYAFAILL	PMDFYIAVCV	EFDLGFSIQM	QFQFFTEHGF	RLVX	
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from *N. gonorrhoeae*:

m018/g018

	10	20	30	40	50	60
m018.pep	MQQRQLVGRV	ACDEDMRNTG	LHGQRVGNRY	AARIFFDIDI	FQTDIVNVRT	AAHGQHIIFG
		:	:	:	:	:
g018	MQQQQLVGRV	ARNKDMRNAG	LHGQRVGNRY	AARVFDIDV	FQTDIVNVRT	ATYGCQHIIFG
	10	20	30	40	50	60

201

	70	80	90	100
m018.pep	NKYAFFAILLPMDFYIAVCIEFDLGFSIQMVFQFFAEHGVRLVX			
	: : :			
g018	NKYAFFAILLPMDFYIAVCVEFDLGFSIQMVFQFFSEHGFRLVX			
	70	80	90	100

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 105>:

```

g019.seq (partial)
1  ..ctgctggcgg ccctggtgct tgccgcgtgt tcttcgACAA ACacacTGCC
51  AGCCGGCAAG ACCCGGCAG ACAATATAGA AActgcCgAC CTTTCGGCAA
101 GCGTTCCAC ccgcCTGCC GAACCGAAG GAAAAACGCT GGCAGATTAC
151 GCGGCTACC CGTCCGCACT GGATGCAGTG AAACAGAACA ACGATGCGGC
201 AGCCGCCGCC TATTGGAAA Acgcaggaga cagCGcgatg gcGAAAatg
251 tccgcaagga gtgGCTGa

```

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

```

g019.pep (partial)
1  ..LLAALVLAAC SSTNTLPAGK TPAADNIETAD LSASVPTRPA EPEGKTLADY
51  GGYPSALDAV KQNNDAASAAA YLENAGDSAM AENVRKEWL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 107>:

```

m019.seq (partial)
1  ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
51  GCTTGCCGCG TGTTCTTCGA CAAACACACT GCCAGCCGGC AAGACCCCGG
101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
151 GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
251 AAAACGCCCG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCACGCAG ACAGTGGACG CTGTTTGAC AGGAATACGC
351 CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
451 AAAGTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGC GCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCTT TGGACTATTA
801 CGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
901 ATCTCGCATA TGCCCCGAAA ACTGCAAAAA AGCCCCGACCT GGTCTACTG
951 GCTGGCACGC AGCCGCGCCG CAACGGGCAA CACGCAAGAG GCGGAAAAAC
1001 TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTATATG GGTGCTGGCA
1051 GGGGAAGAAT TGGGTCGGAA AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
1151 CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
1201 CGTCAGGCTC AGCGGGAATG GCGTTTGGC ACACGCGGCT TTGACGAAGA
1251 CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCAGCC GCAAACTCAA CTACACCTTG
1351 CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTGTCGAG GAAAGCCGCT
1451 TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGACGGTT
1501 ATGCTTGCCA CCGCGCGCGA AATCGCCGGC AAAATCGGTA TGGATGCCCG
1551 ACAACTTTAC ACCGCCGACG GG...

```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

```

m019.pep (partial)
1  MYLPSMKHSL PLLAALVLAA CSSTNTLPAG KTPADNIETA DLSASVPTRP

```



```

51  AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARRQWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAEELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWVVL
251 GHYQSQNLNV PAALDYYGKV ADRRQLTDDQ IEWYARAALR ARRWDELASV
301 ISHMPEKLQK SPTWLYWLAR SRAATGNTQE AEKLYKQAAA TGRNIFYAVLA
351 GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQN SQSAGDAKMR
401 RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL
451 RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADG...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 109>:

a019.seq

```

1  ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGCTGCTGG TGGNCCTGGT
51  GCTTGCCGCG TGTTCTTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
101 CAGACAAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
151 GCCGAACCCG AANGAAAAAC GTNGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCGCGCAG ACAGTGGACG CTGTNTGCAC ANGAATATGC
351 NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCCG AACTGGTCAA AAATACGGGC
451 AAACTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCT TGGACTATTA
801 NGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCNNNTNNGC NNNCGNNGTT NGNANGANNT GGCNNCCGNN
901 ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGNCNNANNT NGNTNNANTG
951 NNTGGCAGCG AGCCGCGCCG CNACGGGCAA CACGCAANAN GCGGANAAC
1001 TNTACAAACA GGCGGCAGCA NCGGGCANGA ATTTTATATG NGTGCTGNCN
1051 GGGGAAGAGT TGGGGCGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAANC AGCGTCCTCC GTATGGCGGA AGACGGCGCG ATTAAGCCGC
1151 CGCTGGTGCT GTTCCGAAAC AGCCGAACCG CCGGCGATGC GAAAATGCGC
1201 CGTCNGGCTC AGGCGGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1251 CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCAGCC GCAAACCTCA CTACACCTTG
1351 CGTACATTT CGNNNNNTNA NGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCGTGGG TTTACGGGCT GATTTCGTCAG GAAAGCCGCT
1451 TCGTTATGGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501 ATGCTTGCCA CCGCGCGCGA AATCGCCGCG AAAATCGGTA TGGATGCCGC
1551 ACAACTTTAC ACCGCCGACG GCAATATCCG TATGGGGACG TGGTATATGG
1601 CGGACACCAA ACGCCGCGCT CAAAACAACG AAGTCCTCGC CACCGCAGGC
1651 TATAACGCCG GTCCCGGCAG GGCGCGCCGA TGGCAGGCGG ACACGCCCCCT
1701 CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTTCCGAA ACGCGCGACT
1751 ATGTCAAAAA AGTGATGGCC AATGCCGCTC ACTACGCCTC CCTCTTCGGC
1801 GCGCCGCACA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCCGCTG
1851 A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

a019.pep

```

1  MYPPSLKHSL PLLVLVLAA CSXTNTLSAD KTPADNIETA DLSASVPTXP
51  AEPEKTXAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARRQWT LXAXEYAKLE PAXRAQEVEC YADSSRNDYT RAEELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWVVL
251 GHYQSQNLNV PAALDYXGKV ADRRQLTDDQ IEWYARAAXX XRXXXXXXXX
301 XXXXXXXXXX XXXXXXXXAR SRAATGNTQX AXKLYKQAAA XGXNIFYAVLX
351 GEELGRXIDT RNNVPDAGKX SVLRMAEDGA IKRALVLFNRN SRTAGDAKMR
401 RXAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL

```

451 RYISXXXDTV IRHAQNVNVD PAWVYGLIRQ ESRFVMAQS RVGAQGLMQV
 501 MPATAREIAG KIGMDAAQLY TADGNIRMGY WYMADTKRRL QNNEVLATAG
 551 YNAGPGRARR WQADTFLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG
 601 APHIPLKQRM GIVPAR*

m019/a019 88.9% identity over a 524 aa overlap

m019.pep	10	20	30	40	50	60
	MYLP	SMKH	SLPL	LAAL	VLAACS	STNTLP
a019	10	20	30	40	50	60
	MYPP	SLKH	SLPL	LVXL	VLAACS	XTNTLS
m019.pep	70	80	90	100	110	120
	YGGY	PSAL	DAVK	QKND	AAVAAY	LENAGD
a019	70	80	90	100	110	120
	YGGY	PSAL	DAVK	QKND	AAVAAY	LENAGD
m019.pep	130	140	150	160	170	180
	PAGRA	QEVE	CYAD	SSRN	DYTRAA	ELVKNT
a019	130	140	150	160	170	180
	PAXRA	QEVE	CYAD	SSRN	DYTRAA	ELVKNT
m019.pep	190	200	210	220	230	240
	LLAGR	QTTD	ARNL	AAAL	GSFP	DGGTQ
a019	190	200	210	220	230	240
	LLAGR	QTTD	ARNL	AAAL	GSFP	DGGTQ
m019.pep	250	260	270	280	290	300
	EQRSF	AWGV	LGHY	QSQNL	NVPAAL	DYVGK
a019	250	260	270	280	290	300
	EQRSF	AWGV	LGHY	QSQNL	NVPAAL	DYVGK
m019.pep	310	320	330	340	350	360
	ISHMP	EKLQ	KSPTW	LYWL	ARSRA	TGNTQ
a019	310	320	330	340	350	360
	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
m019.pep	370	380	390	400	410	420
	RNNVP	DAGK	NSVRR	MAED	GAVK	RALVLF
a019	370	380	390	400	410	420
	RNNVP	DAGK	XSVL	RMAED	GAIK	RALVLF
m019.pep	430	440	450	460	470	480
	TAAQT	AFDH	GFYD	MAVNS	AE	TRDK
a019	430	440	450	460	470	480
	TAAQT	AFDH	GFYD	MAVNS	AE	TRDK
m019.pep	490	500	510	520		
	ESRFV	IGAQS	RVGA	QGLMQ	VMPAT	AREIAG
a019	490	500	510	520	530	540
	ESRFV	IGAQS	RVGA	QGLMQ	VMPAT	AREIAG
a019	550	560	570	580	590	600
	QNNE	VLAT	AGYN	AGPGR	ARRW	QADTF

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from *N. gonorrhoeae*:

g019/m019

		10	20	30	40	49
g019.pep		LLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD				
m019	MYLP	SMKHSLPL	LAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAE	PERKTLAD		
		10	20	30	40	50 60
	50	60	70	80	89	
g019.pep	YGGYPSALDAVKQNNDAAAAYLENAGDSAMAENVRKEWL					
m019	YGGYPSALDAVKQNDAAVAAYLENAGDSAMAENVRNEWLKS LGARRQWTLFAQEYAKLE					
		70	80	90	100	110 120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 111>:

g023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  AATGCAGCGT GCGACTGCGG TTATTATGTT GATTTATAAC GTTGCACTTT
101 TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
151 TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
201 ATTCTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTCGG CGTGCGTTT TTTTTCAGG TTGCCACCAT TGtctGGCTG
301 GTCGGCTGCC TCGTGATTTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>:

g023.pep

```

1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFA PKEYPAWQAF
51  FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

m023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  GATGCAACGT GCGACTGCGG TTATTATGTT GATTTATAAC GTTGCACTTT
101 TAGTGGTTCT ATTTCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
201 ATTCTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTCGG CGTGCGTTT TTTTTCAGG TTGCCACCAT CGTTTGGCTG
301 GTCGGCTGTC TCGTGATTTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>:

m023.pep

```

1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFS PKEYSAWQAF
51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

a023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GGGATTGGGC
51  GATGCAACGT GCGACCGCGG TTATTATGTT GATTTATAAC GTTGCACTTT
101 TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT

```

```

201  ATCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATNA
251  AACCTTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT CGTCTGGCTG
301  GTCGGCTGCT TGGTGTATTC AATTAAAGTA ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

```

a023.pep
1   MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFA L PKEYSAWQAF
51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYXKPGVRL FLQVATIVWL
101 VGCLVYSIKV IWG*

```

m023/a023 96.5% identity over a 113 aa overlap

```

          10      20      30      40      50      60
m023.pep  MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFS L PKEYSAWQAF FSQTWVKVFT
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a023      MVERKLTGAHYGLRDWAMQRATAVIMLIYT VALLVVLFA L PKEYSAWQAF FSQTWVKVFT
          10      20      30      40      50      60

          70      80      90      100     110
m023.pep  QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a023      QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSIKVIWGX
          70      80      90      100     110

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng) from *N. gonorrhoeae*:

```

g023/m023
          10      20      30      40      50      60
g023.pep  MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFA L PKEYPAWQAF FSQAWVKVFT
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m023      MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFS L PKEYSAWQAF FSQTWVKVFT
          10      20      30      40      50      60

          70      80      90      100     110
g023.pep  QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m023      QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX
          70      80      90      100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 117>:

```

g025.seq
1   ATGTTGAAAC AAACgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
51  GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
151 ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
201 CGTGCAAact gcgcgggttT ATTGCGCTCC TGCTTATGTT CCGCcgTCTG
251 CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
301 aacgCGCGa cgCataCTAT TGTGCGTGGC GACACgGtgt acaACATTtC
351 caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCCTGG AACGGCATGA
401 CCGACAATAC GTTGAACATC GGTGAGATTG TTAAAGTCAA ACCGGCagga
451 TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
501 TGCCGCGCAA ACCCCTGTGA AACC CGCCGc gcaACCGCCC GTTCAGTCCG
551 CGCCGCAACC TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
601 GCGCCGCCCG CGCAATCTCC TGCCGCTTCG CCTTCCGGCA CGCGTTCGGT
651 CGGCGGCATT GTTTGGCAGC GTCCGACCCA AGGTAAAGTG GTTCCGATT

```

```

701 TCGGCGGCGG CAACAAGGGT GTCGATATTG CCGGCAATGC CGGACAACCC
751 GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGGTT CAGGTTTGAG
801 GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
851 CGTACGGGCA CAACCAAAAA TTGCTGGTCG GCGAAGGTCA GCAGGTCAAA
901 CGCGGTCAGC AGGTTGCTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
951 GCTTCATTTT GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
1001 ATATCGCGTT CTGA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>:

```

g025.pep
  1 MLKQTTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
 51 TPNNAAPAAAN DAPYVPPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
101 NAATHTIVRG DTVYNISKRY HISQDDFRW NGMTDNTLSI GQIVKVKPAG
151 YAAPKTAAVE SRPAVPAAQ TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
201 APAPQSPAAS PSGTRSVGGI VWQRPTQGV VADFGGNGK VDIAGNAGQP
251 VLAAADGKVV YAGSGLRGGY NLVIIQHNS FLTAYGHNQK LLVGEGQQVK
301 RGQOVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```

m025.seq (partial)
  1 ..GTGCCGCGG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
 51 GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
101 TCGACATCAA CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
151 AACATTTCCA AACGCTACCA TATCTCTCAA GACGATTTC GTGCGTGGAA
201 CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
251 CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
301 GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
351 CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
401 AACGCTACCA TATCTCTCAA GACGATTTC GTGCGTGGAA CGGCATGACC
451 GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
501 TGCCGCACCG AAAACCCGAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
551 CCGTGCAAAC CCCTGTGAAA CCCGCCGCGC AACCGCCTGT GCAGTCCGCG
601 CCGCAACCTG CCGCGCCCGC TGCGGAAAT AAAGCGGTT CCGCGCCCGC
651 CCGCAATCT CCTGCCGCTT CGCCTTCCGG CACGCGTTCT GTCGGCGGCA
701 TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGGTTGCCGA TTTCGGCGGC
751 AACAACAAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTGGC
801 GGCGGCTGAC GGCAAAGTGG TTTATGCCGG TTCAGGTTT AGGGGATACG
851 GAAACTTGGT CATCATCCAG CATAATTCTT CTTTCCTGAC CGCATACGGG
901 CACAACCAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACCGGTCA
951 GCAGGTTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGCTTCATT
1001 TCGAGGTGCG TCAAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051 TTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

```

m025.pep (partial)
  1 ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
 51 NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAVKSIPA
101 VPAAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
151 DNMLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
201 PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVADFGG
251 NNKGVDIAGN AGQPVLAAD GKVVYAGSGL RGYGNLVIIQ HNSSLTAYG
301 HNQKLLVGEG QQVKRGQQA LMGNTDASRT QLHFVVRQNG KPVNPNSYIA
351 F*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

a025.seq
  1 ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
 51 GTTGGGCGGA TGCCCCACCC AACACCCTT TCCTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTACCG TCTGCGCCG TTTACAATCC TTATGGCGCA

```

```
151 ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCCGCC
201 GGTGCAAAGC GCGCCGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
251 CACCTGCCGT TTCGGGTACA TACGTTCTT CTTACGCANC CGTCGACATC
301 AACGCGGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTC
351 CAAATGCTAC CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401 CCGACAATAC GTTGAAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA
451 TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
501 TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACGCGGCGA
551 CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTC CAAACGCTAC
601 CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA CCGACAATAC
651 GTTGAAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
701 CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTCGCA
751 ACCCCTGTGA AACCCGCCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
801 TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCGCCC GCCCGCAAT
851 CTCCTGCCGC TTCGCTTCC GGCACGCGTT CCGTCGGCGG CATTGTTTGG
901 CAGCGTCCGA CGCAAGGTAA AGTGTTGCC GATTTCGGCG GCAACAACAA
951 GGTGTTCGAT ATTGCAGGAA ATGCGGGACA GCCCGTTTTG GCGGCGGCTG
1001 ACGGCAAAAGT GGTATTATGCA GGTTCGGTT TGAGGGGATA CGGCAATTTG
1051 GTCATCATCC AGCATAATTC TTCCTTCTG ACCGCATACG GGCACAACCA
1101 AAAATTGCTG GTCGGCGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTG
1151 CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGTTCA TTTGAGGTG
1201 CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 025.a>:

```
a025.pep
  1 MLTP TTL*VA CTALAAQLGG CPTQHPSPIV AGNSGMQTV SAPVYNPYGA
 51 TPNAAAPAAN DAPYVPPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
101 NAATH TIVRG DTVYKISKY HISQDD FRAW NGMTDN TLSI GQIVKVPAG
151 YAAPKAAAVK SRPAVPAAAQ PLVQSAPVDI NAATH TIVRG DTVYNISKRY
201 HISQDD FRAW NGMTDN TLSI GQIVKVPAG YAAPKAAAVK SRPAVPAAVQ
251 TPVKPAAQPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
301 QRPTQ GKVVA DFGGNNKGVD IAGNAGQPV LAAADGKVYA GSGLRGYGNL
351 VIIQHNS SFL TAYGHNQKLL VEGEQQVKRG QQVALMGNT EASRTQLHFEV
401 RQNGKPVNPN SYIAF*

m025/a025 97.4% identity over a 351 aa overlap

                                10      20      30
m025.pep                                VPPVQSAPVYTTPPAYVPPSAPAVSGTYVPS
                                |||||:|||||
a025      GMQTVPSAPVYNPYGATPNAAAPAANDAPYVPPVQSAPVYXPPAYVPPSAPAVSGTYVPS
              40      50      60      70      80      90

              40      50      60      70      80      90
m025.pep      YAPVDINAATH TIVRGDTVYNISKRYHISQDD FRAWNGMTDN TLSIGQIVKVPAGYAAP
              || |||||:|||||
a025      YAXVDINAATH TIVRGDTVYKISKYHISQDD FRAWNGMTDN TLSIGQIVKVPAGYAAP
              100     110     120     130     140     150

              100     110     120     130     140     150
m025.pep      KAAAVKSRPAVPAAAQPPVQSAPVDINAATH TIVRGDTVYNISKRYHISQDD FRAWNGMT
              |||||
a025      KAAAVKSRPAVPAAAQPLVQSAPVDINAATH TIVRGDTVYNISKRYHISQDD FRAWNGMT
              160     170     180     190     200     210

              160     170     180     190     200     210
m025.pep      DNMLSIGQIVKVPAGYAAPKTA AVESRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
              || |||||:|||||
a025      DNTLSIGQIVKVPAGYAAPKAAAVKSRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
              220     230     240     250     260     270

              220     230     240     250     260     270
m025.pep      KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQ GKVVADFGGNNKGVDIAGNAGQPVLAAD
              |||||
```

208

```

a025      KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
           280      290      300      310      320      330

m025.pep      280      290      300      310      320      330
GKV VYAGSGLRGYGNLVIIQHNS SFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDASRT
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a025      GKV VYAGSGLRGYGNLVIIQHNS SFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTEASRT
           340      350      360      370      380      390

m025.pep      340      350
QLHFEVRQNGKPVNPNSYIAFX
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a025      QLHFEVRQNGKPVNPNSYIAFX
           400      410

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from *N. gonorrhoeae*:

m025/g025

```

m025.pep      10      20      30
              VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g025      GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS
           40      50      60      70      80      90

m025.pep      40      50      60      70      80      90
YAPVDINAATH TIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVPAGYAAP
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g025      YAPVDINAATH TIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVPAGYAAP
           100     110     120     130     140     150

m025.pep      100     110     120     130     140     150
KAAAVKSRPAVPAAAQPPVQSAPVDINAATH TIVRGDTVYNISKRYHISQDDFRAWNGMT
|
g025      K-----

m025.pep      160     170     180     190     200     210
DNMLSIGQIVKVPAGYAAPKTAAVESRPAPVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g025      -----TAAVESRPAPVPAVAAAQTPVKPAAQPPVQSAPQPAAPAAEN
           160     170     180     190

m025.pep      220     230     240     250     260
KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g025      KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
           200     210     220     230     240     250

m025.pep      270     280     290     300     310     320
ADGKV VYAGSGLRGYGNLVIIQHNS SFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g025      ADGKV VYAGSGLRGYGNLVIIQHNS SFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
           260     270     280     290     300     310

m025.pep      330     340     350
RTQLHFEVRQNGKPVNPNSYIAFX

```

209

g025 |||||
 RTQLHFEVRQNGKPVNPNSYIAFX
 320 330

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 113>:

g031.seq

```

1  ATGGGTGTCCC TCCGCTTCAG ATTCGGCAAC CACTTTAAAC GCCGACATTC
51  TGACAATTTT CTTTTCCGCC AGCCAAATAT CATGCGTATC TTTCGGTTTCG
101 GGCTTGTGTT GCATGGCAAC CTTCAACAGC CGCGCCATCA CAGGAATCGT
151 CGTTCCCTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA
201 ACAGCAGGTG CGAATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC
251 GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC
301 ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC
351 GCGCGACCAG CATCAGGAAC GCCGCAATCG CCAAGGCTTC GCCTGCCCTG
401 TCCAACACGC CGGCGGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT
451 TGCCTGCGCC AGCCAAGCCA AACCCTCCAT CACACGCAA ACCTGTTCCG
501 TgcACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC
551 AAAAAAGCCG TGCCGCCTAT GGTATTGGTA AACGCAAACA CAAGCAGCCC
601 GCCCGACACA ATCATCAGCG CGTACAGACC TTCCGtacac acctccaatt
651 cccaatcaac gtcatactgt tctcccgtgt taaaatgttc ttcacttcag
701 aatccccccc ttcttcccag cccgaaacct tcatgtgtta naccctgggg
751 tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgccccctt
801 cgcccgcctt ctccttcggt gaaaacttgt tgtcccgtc ttacattaa
  
```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>:

g031.pep

```

1  MVSLRFRFGN HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR
51  RSLNQQRQHH HGKRHIKQOV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF
101 TQAVIEFPQT AEHCQRTRDQ HQERRNRQGF RRPVQHAGGR NQQTEDHEQS
151 CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAAY GIGKRKHQKP
201 ARHNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPSSQ PETFMCXTLG
251 CPNGFSNLPM TLLVAPFARE LLPGKLVVPV LH*
  
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

m031.seq (partial)

```

1  ...CGCCTGAAGC ACGGTGTCCG ACTGCATTTC TATTCGGCTA TACGCCTTTT
51  CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA
101 CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAGGCTT CCGCCGCCCT
151 GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAAACGTG
201 CGGACAAAGT GGTGCAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC
251 AG....
  
```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>:

m031.pep (partial)

```

1  ...RLKHGVLHF YSAIRLFTQA VIEFPQTAEH CRRTRDQHQE RRRNRQGFRRP
51  VQHVGRNRNQ QRHSQTCGQS GRNHAQKQQC ATRQ....
  
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

a031.seq

```

1  ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA
51  CTGCCGCGCG ACGCGCGACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT
101 TCCGCCGCCG CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT
151 TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAACAGTG
201 CGCCACCCGG CAG
  
```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

a031.pep (partial)

```

1  IRLFTQAVIE FPQTAECRR TRDQHQERRN RQGFRFPVQH VGRNRQQQRH
51  SQTGQSGRN HAQKQQCATR Q
  
```


210

m031/a031 100.0% identity over a 71 aa overlap

	10	20	30	40	50	60
m031.pep	RLKHGVLH	FYSAIRLFTQ	AVIEFPQTAEH	CCRTRDQH	QERRNRQG	FRRPVQHVGRNRQQ
a031		IRLFTQAVIEFPQTAEH	CCRTRDQH	QERRNRQG	FRRPVQHVGRNRQQ	
		10	20	30	40	

	70	80
m031.pep	QRHSQTCGQSGRNHAQKQQCATRQ	
a031	QRHSQTCGQSGRNHAQKQQCATRQ	
	50	60

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF 031.ng) from *N. gonorrhoeae*:

m031/g031

		10	20	30
m031.pep		RLKHGVLH	FYSAIRLFTQ	AVIEFPQTAEH
g031	NQQRQH	HHGKRHIKQQV	RIGNAHHRQ	RNRNRYGSSQAQPTDIRLFTQAVIEFPQTAEH
	60	70	80	90

	40	50	60	70	80
m031.pep	CCRTRDQH	QERRNRQG	FRRPVQHVGRNRQQ	QRHS-QTCGQSGRNHAQKQQCATRQ	
g031	CQRTRDQH	QERRNRQG	FRRPVQHVGRNRQQ	TEHDEQSLRQPSQTVHHTQNVFRRTVALV	
	120	130	140	150	160

g031	TDNDAGKVN	RQKAAAAYG	IGKRKHQPAR	HNHQRVQTF	FRTHLQFPIN	VIASRVKMF	FTS
	180	190	200	210	220	230	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 119>:

g032.seq

```

1  ATGCGGCGAA ACGTGCCTGC CGTCGCCGTA TTGCGCCGCC CACGATTCGA
51  GGCGTTTTTG GATTTGCGGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAAGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGACGC TGCTTGCGCC
201 CTTTGCCGGT AACGTGTACC CACGCTTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGCTC
301 GAACAGCGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAACAT CCGGTTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGCGG GTCGGGTCGG CGCGCATCAG
451 CCCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCACGCC GACAGCTTGC
501 GCGCCAGCGT CCGACCGTCC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551 GccgTAAAT CGCGCCGCGA CAAGTCCTGC GGCACGCcgc ctgcaTCTTC
601 AGACGGCATT TGTGCCAACA GTGCAAACAG TTCTTCCAA TCGCGCCGGT
651 ATGCCGCAAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAA
701 TAAGCGTCAA AATacgcgc AAACccgTCC AAAACCATAA CCGTCCCACA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 032.ng>:

g032.pep

```

1  MRRNVPAAV LRRPRFEAF DLALAARAV PAGKQGFVR CRLTQRQIVF
51  QGFHAFAGR NLTLAPFAG NVYPRFVQIY IICIQAVYLA HAQTAHVHQL

```

```

101 EQRVVAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGAHQ
151 PAFDQPGAIL PPRRQLARQR PTVQTALRQP PQRRRKIAPR QVLRHAACIF
201 RRHLCQQCKQ FFQIAPVCRN RVLRLALAHQ VFQISVKIRR KPVQNHNRPT
251 QISKNQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

m032.seq (partial)
  1 ATGCGGCGAA ACGTGCMTGC mGTCGCCGTT KTGCGCCGCC CATTGCGCCA
  51 AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
201 CTTTGCCGAT AAcGTGTACC CACGCyTCGT CCAAATAGAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
301 GAACAGGGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTcAGC CATTcCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGCGG GTCGGGTCGG CGTGCATCAG
451 GCCGCGCTTT ACCAGCCAAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
501 GAGCCAGCGT CCGTCCCCC AAACCGCG...

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:

```

m032.pep (partial)
  1 MRRNVXAVAV XRRPLRQTFE DLALAQARAV PAGKQGFVAV CRLTORQIVF
  51 QGFHAFADQR HLPLXAPFAD NVYPRXVQID IICIQAVYLA HAQTAAVHQF
101 EQGVVAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGVHQ
151 AALYQPNAIL PPRRKLASQR PEPQTA...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```

a032.seq
  1 ATGCGGCGAA ACGTGCTGCG CGTCGCCGTT TTGCGCCGCC CATTGCGCCA
  51 AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCGTC
201 CTTTGCCGGT AACGTGTACC CACGCCTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
301 GAACAGCGCG TGATCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTcAGC CATTcCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGCGG GTCGGGTCGG TATGCAGCAG
451 ACCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
501 GCGCCAGCGT CCGCGCATTC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551 GCCGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
601 AGACGGCATT TGTGCAACA GCGCAAACAG TTCTTCCAAA TCGCGCCGGT
651 ATGCCGCCAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATGCGCCGC AAACCCGTCC AAAACCATAA CCGCCCCACA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:

```

a032.pep
  1 MRRNVPAAVAV LRRPLRQTFE DLALAQARAV PAGKQGFVAV CRLTORQIVF
  51 QGFHAFAGQR NLPLASFAG NVYPRLVQIY IICIQAVYLA HAQTAAVHQF
101 EQRVIAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGMQQ
151 TAFDQPGAIL PPRRQLARQR PRIQTALRQP PQRRRKIALR QALRHAACIF
201 RRHLCQQRKQ FFQIAPVCRH RVLRLALAHQ VFQISVKMRR KPVQNHNRPT
251 QISKKQ*

```

m032/a032 88.1% identity over a 176 aa overlap

```

m032.pep      10      20      30      40      50      60
MRRNVXAVAVXRRPLRQTFDLALAQARAVPAGKQGFVAVRCRLTORQIVFQGFHAFADQR
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a032          10      20      30      40      50      60
MRRNVPAAVAVLRRPLRQTFDLALAQARAVPAGKQGFVAVRCRLTORQIVFQGFHAFAGQR

```

Homology with a predicted ORF from *N. gonorrhoeae*

m032/q032

	10	20	30	40	50	60
m032.pep	MRRNVXAVAVXRRPLRQTF	LDLALAQA	RAVPAGKQGF	AVRCRLTQRQ	IVFQGF	HAFADQR
g032	MRRNVPAVAVLRRPRFEA	FLDLALAQA	RAVPAGKQGF	AVRCRLTQRQ	IVFQGF	HAFAGQR
	10	20	30	40	50	60
	70	80	90	100	110	120
m032.pep	HLPLXAPFADNVYPRXVQ	IDIIICIAVYLA	HAQTA	AAVHQFEQ	GVVAHRQR	VAAVHGQIQH
	:					
g032	NLTLLAPFAGNVYPRFVQ	IYIIICIAVYLA	HAQTA	AAVHQLEQ	RVVVAHRQR	VAAVHGQIQH
	70	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPF	FLRQGF	YALGLL	RRFDV	GGRVGV	HQAALYQPNAILP
g032	PVQPF	FLRQGF	YALGLL	RRFDV	GGRVGA	HQAPAFDQPGAILP
	130	140	150	160	170	180
g032	PQRRRK	IAPRQV	LRHAAC	I	FRRHLC	QQCKQFFQIAPVC
	190	200	210	220	230	240

g033.seq

1	ATGGCGGCGG	CGGACAAACT	CTTGGGCGGC	GACCGCCGCA	GCGTCGCCAT
51	CATCGAGAC	GGCGCGATG	CGGCGGGGA	GGCGTTTGA	GCCTTGAATT
101	CGCGGGGCA	TATGGATGT	GATTTGCTGG	TCGCTACTCA	CGACAACGAA
151	ATGTGCAATT	CCCCAACGT	CGGCGGTTG	CCCAATATC	TTGCCAGCAA
201	CGTCGTGCGC	GATATGCACG	GACTGTTGAG	TACCGTCAA	GCGCAAAcgg
251	GCAAGGTATT	AGACAAAATA	CCCGGCGCGA	TGGagtTTGC	CCAAAAAGTC
301	GAACCaAA	TCAAAACCT	TGCCGAAGAA	GCCGAACACG	CCAAACAGTC
351	GCTGTCGCT	TTTGAAAAAT	TCGGCTTCG	CTACACCGG	CCCGTGGACG
401	GACACAAGT	CGAGAAATCT	TGGAGCTAT	TGAAAGACTT	GCGCAGCCGC
451	AAAGGCCCTC	AGTTGCTGCA	CGTCATCACC	AAAAAGGGCA	ACGCTACAA
501	ACTCGCCGAA	AACGACCCcg	tcaAATACCA	CGCCGTCGc	aACCTGCcta
551	AAGAAGGCGG	GGCGCAATg	ccGTCGTGAA	AAGAACCCAA	GCCCGCCgCc
601	aaaccgACCT	ATACCCAAGT	ATTCCGCAAA	TGGCTGTGCG	ACCGGGCGGC
651	GGCAGATTCC	CGACTGGTTG	CGATTACCCC	CGCATGCGC	GAGGGCAGCG

```

701 GACTGGTGGG GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT tacCTTTGCC GGCGGTTTGG CGTGCGAAGG
801 CATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCCG TTTGTTTGCC
901 GTCGACCGTG CGGGCATCGT CGGCGCGGAC GGTCCGACCC ATGCCGGCTT
951 GTACGATTTG AGCTTCTTGC GCTGTGTGCC GAACATGATT GTTGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCG
1051 GATGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTTCAGAC GGCATGAAAA CCGTGGAAT CCGCAAGGCG ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCC TTcatTGCCT TCGGCAGTAT GGTGCCCACC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTt
1251 cgtcaaacCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCAcg
1301 accGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGCG
1351 GCGGTCTTGG AAGTGTGGC GAAACACGGC ATCTGCAAAC CCGTTTGTCT
1401 TTTGGGCGTT GCCGATACCG TAACCGAACA CGGCGATCCG AAAAAACTTT
1451 TGGACGATTT GGGTTTGAGT GCCGAAGCGG TGAACGCCG GGTGCGCGAG
1501 TGGCTGCCCG ACCGTGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>:

g033.pep

```

1 MAAADKLLGG DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAE AEHAKQSLSL FENFGFRYT PVDGHNVENL VDLKDLRSR
151 KGPQLLHVIT KKGNGYKLA NDVKYHAVA NLPKEGGAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAT
401 ALAVAEKLNA TVADMRFPVK IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLLG ADTVTEHGDP KKLLDDLGLS AEAVERRVRE
501 WLPDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

m033.seq

```

1 ATGGCGGCGG CAGACAAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGGCGAC GGCGCGATGA CGGCGGGGCA GGCGTTTGAA GCCTTGAATT
101 GCGCaG.CGA TATGGATGTr GATTTGCTrG TCGTCCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAACGT CGGCGCGCTG CCGAAATACC TTGCCAGCAA
201 GTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAA GCGCAACCGG
251 CCAAGGTATT AGACAAAATA CCGGCGCGCA TGGAGTTTGC CCAAAAGTGC
301 GAACACAAAA TCAAACCCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
351 GCTGTCTTTG TTTGAAAAC TCGGCTTCCG CTACACCGGC CCCGTGGACG
401 GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451 AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTCGCC AACCTGCCA
551 AAGAAAAGCG GGCGCAAATG CCGTCTGAAA AAGAACCCTA GCCCGCCGCC
601 AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC
901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCCAGCC ATGCCGGTCT
951 GTACGATTTG AGCTTTTTGC GCTGCGTGCC GAACATGATT GTCGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTCGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTTCAGAC GGCATGAAAA CCGTGGAAT CCGCAAGGCG ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTGCGCCCC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTGCGCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG

```

```

1301 ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTGTCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>:

```

m033.pep
  1 MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
  51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
 101 EHKKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLDLRSR
 151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
 201 KPTYTQVFGK WLCDRRAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
 251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
 301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSCTCYQA
 351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
 401 ALAVAEKLNA TVADMRVVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
 451 AVLEVLAHKG ICKPVLLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
 501 WLSDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```

a033.seq
  1 ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
  51 CATCGGCGAC GCGCGATGA CCGCGGGTCA GCGGTTTGAA GCCTTGAAC
 101 GCGCGGGCGA TATGGATGTG GATTTGCTGG TCGTCCTCAA CGACAACGAA
 151 ATGTCGATTT CCCCACAGT CCGTGCGTTG CCCAAATACC TTGCCAGCAA
 201 CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAAA GCGCAAACGG
 251 GCAAGGTATT AGACAAAATA CCCGCGCGCA TGGAGTTTGC CCAAAAAGTC
 301 GAACATAAAA TCAAACCCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
 351 ACTGTCTTTG TTTGAAAACT TCGGCTTCCG CTATACCGGC CCCGTGGACG
 401 GACACAACGT CGAAAATCTG GTCGATGTAT TGGAAAGACCT GCGCGGACGC
 451 AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
 501 ACTCGCCGAA AACGATCCCG TCAAATACCA CGCCGTCGCC AACCTGCCA
 551 AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCAA GCCCGCCGCC
 601 AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
 651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
 701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751 ATCGCCGAGC AGCAGCCCGT TACCTTTGCC GCGGTTTGG CTTGCGAAGG
 801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
 901 GTCGACCGCG CGGGCATCGT CCGCGCGGAC GGCCCGACCC ATGCCGGTTT
 951 GTACGATTTA AGCTTTTTGC GCTGCATTCC GAATATGATT GTCGCCGCGC
1001 CGAGCGATGA AAATGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGTGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAT CCGCAAGGGC ATTATCCGCC
1151 CGGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCCT
1201 GCATTGGCGG TCGCCGAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
1301 ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGAG
1351 GCGGTGCTGG AAGTGTGGC GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

```

a033.pep
  1 MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
  51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
 101 EHKKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLEDLRGR
 151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
 201 KPTYTQVFGK WLCDRRAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
 251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA

```

```

301 VDRAGIVGAD GPTHAGLYDL SFLRCIPNMI VAAPSDENEC RLLSTCYQA
351 DAPAAVRYPR GTGTGVPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401 ALAVAGKLNA TVADMRFBVP IDEELIVRLA RSHDRIVTLE ENAEQGGAGS
451 AVLEVLAKHG ICKPVLLLG VADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501 WLSDRDAAN*

```

m033/a033 98.4% identity over a 509 aa overlap

m033.pep	10	20	30	40	50	60
	MAAADKLLGSDRRSVAIIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISP NVGAL					
a033	MAAADKQLGSDRRSVAIIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISP NVGAL					
	10	20	30	40	50	60
m033.pep	70	80	90	100	110	120
	PKYLASNVRDMHGLLSTVKAQTGKVLDPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
a033	PKYLASNVRDMHGLLSTVKAQTGKVLDPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
	70	80	90	100	110	120
m033.pep	130	140	150	160	170	180
	FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPKQLLHVITKKGNGYKLAENDPVKYHAVA					
a033	FENFGFRYTGPVDGHNVENLVDVLEDLRGRKGPKQLLHVITKKGNGYKLAENDPVKYHAVA					
	130	140	150	160	170	180
m033.pep	190	200	210	220	230	240
	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ					
a033	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ					
	190	200	210	220	230	240
m033.pep	250	260	270	280	290	300
	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA					
a033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA					
	250	260	270	280	290	300
m033.pep	310	320	330	340	350	360
	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAAPSDENECRLLSTCYQADAPAAVRYPR					
a033	VDRAGIVGADGPTHAGLYDLSFLRCIPNMIVAAAPSDENECRLLSTCYQADAPAAVRYPR					
	310	320	330	340	350	360
m033.pep	370	380	390	400	410	420
	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFFIAFGSMVAPALAVAELNATVADMRFBVP					
a033	GTGTGVPVSDGMETVEIGKGIIRREGEKTAFFIAFGSMVAPALAVAGKLNATVADMRFBVP					
	370	380	390	400	410	420
m033.pep	430	440	450	460	470	480
	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLG VADTVTGHGDP					
a033	IDEELIVRLARSHDRIVTLEENAEQGGAGSAVLEVLAKHGICKPVLLLG VADTVTGHGDP					
	430	440	450	460	470	480
m033.pep	490	500	510			
	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
a033	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
	490	500	510			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from *N. gonorrhoeae*:

m033/g033

m033 . pep	MAAADKLLGSDRRSVAIIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISP NVGAL	60
g033	MAAADKLLGGDRRSVAIIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISP NVGAL	60
m033 . pep	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEHAKQSLSL	120
g033	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEHAKQSLSL	120
m033 . pep	FENFGFRYTGPVDGHNVENLVDVLKDLRSRKG PQLLHVITKKNGYKLAENDPVKYHAVA	180
g033	FENFGFRYTGPVDGHNVENLVDVLKDLRSRKG PQLLHVITKKNGYKLAENDPVKYHAVA	180
m033 . pep	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSR LVAITPAMREGSGLVEFEQ	240
g033	NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSR LVAITPAMREGSGLVEFEQ	240
m033 . pep	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNL PVLFA	300
g033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNL PVLFA	300
m033 . pep	VDRAGIVGADGPTHAGLYDLSFLRCVPMIIVAAPS DENECRLLLSTCYQADAPAAVRYPR	360
g033	VDRAGIVGADGPTHAGLYDLSFLRCVPMIIVAAPS DENECRLLLSTCYQADAPAAVRYPR	360
m033 . pep	GTGTGAPVSDGMETVEIGKGIIRREG EKTAFI AFGSMVAPALAVA EKL NATVADMR FVKP	420
g033	GTGTGAPVSDGMETVEIGKGIIRREG EKTAFI AFGSMVAT ALAVA EKL NATVADMR FVKP	420
m033 . pep	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP	480
g033	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTEHGDP	480
m033 . pep	KKLLDDLGLSAEAVERRRVRAWLSDRDAANX	510
g033	KKLLDDLGLSAEAVERRRVREWLPRDAANX	510

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 121>:

g034 . seq

1	ATGAGCCGTT	TATGGTTTTT	TGCCGTAAAA	AACATTATAA	TCCGCCTTAT
51	TTACCTATTG	CCCAAGGAGA	CACAAATGGC	ACTCGTATCC	ATGCGCCAAC
101	TGCTTGACCA	CGCCGCCGAA	AACAGCTACG	GCCTGCCCCG	GTTCAACGTC
151	AACAACCTCG	AACAAATGCG	CGCCATTATG	GAAGCCGCCG	ACCAAGTCAA
201	CGCGCCCGTC	ATCGTACAGG	CGAGCGCAGG	TGCGCGCAAA	TACGcggGCG
251	CGCCGTTTTT	GCGCCACCTG	ATTCTGGCGG	CAGTCGAAGA	ATTTCCGCAC
301	ATCCCCGTCTG	TGATGCACCA	AGACCACGGC	GCATCGCCCC	ACGTgtgCCA
351	ACGCTCCATC	CAACTGGGCT	TCTCCTCCGT	GATGATGGAC	GGCTCTTTGC
401	TCGAAGACGG	CAAAACCCCT	TCTTCTTACG	AATACAACGT	CAACGCCACC
451	CGTACCGTCG	TCAACTTCTC	CCACGCCTGC	GGCGTGTCCG	TCGAAGGCGA
501	AATCGGCGTA	TTGGGCAACC	TCGAAACCGG	CGAAGCAGGC	GAAGAAGACG
551	GAGTGGGCGC	GGCAGGCAAA	CTCTCACACG	ACCAAATGCT	CACCAGCGTT
601	GAAGATGCCG	TGCGTTTCGT	TAAAGATACC	GGCGTTGACG	CATTGGCGAT
651	TGCCGTCTGC	ACCAGCCACG	GCGCATACAA	ATTCACCCGT	CCGCCCCACG

```

701 GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGgctCCAGC TCCGTTCCGC AAGAatgGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
901 ATCGATACCG ACCTGCGCCT CGCTTCCACC GCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>:

```

g034.pep
  1 MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
 51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQRSI QLGFSVMMD GSLLEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAAGK LSHDQMLTSV
201 EDAVRVFKDT GVDALAI AVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV I NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMASRYA KGELNQIVK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 123>:

```

m034.seq (partial)
  1 ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
 51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA WACAGCTACG GCyTGCCGGC GTTCAACGTC
151 AACAACTCG WACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
201 CGCCCCCGTC ATCGTACAGG CGAGTGCCGG TGCGCGCAAA TATGCGGGTG
251 CGCCGTTTTT ACGCCACCTG ATTTTGGCGG CTGTGCAAGT ATTTCCACAC
301 ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
451 CGTACCGTGG TTAACCTCTC CCACGCTTGC GCGTATCCG TTGAAGCCGA
501 AATCGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT GACCAGCGTC
601 GAAGATGCCG TATGTTTCGT TAAAGATACC GCGGTTGACG CATTGGCTAT
651 TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCCACAG
701 GCGATGTATT ACGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA C...

```

This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>:

```

m034.pep (partial)
  1 MSCLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE XSYGLPAFNV
 51 NNLXQMRAIM EAADQVDAPV IVQASAGARK YAGAPFLRHL ILAAVEVFP
101 IPVVMHQDHG ASPDVCQRSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGDAG EEDGVGAVGK LSHDQMLTSV
201 EDAVCFVKDT GVDALAI AVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMH...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 125>:

```

a034.seq
  1 ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
 51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
151 AACAACTCG AACAAATGCG CGCCATTATG GAAGCCGCCG ACCAAGTCAA
201 CGCCCCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAAA TACGCGGGCG
251 CGCCGTTTTT GCGCCACCTG ATTTTGGCGG CTGTGCAAGA ATTTCCGCAC
301 ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TTCTCTCCGT GATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACC

```



```

451 CGTACCGTGG TTAATTTCTC CCACGCCTGC GCGGTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAAACTGG CGAAGCCGCG GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT CACCAGCGTC
601 GAAGATGCCG TGC GTTTCGT TAAAGATACC GCGGTGACG CATTGGCGAT
651 TGCCGTCGGC ACCAGCCACG GCGCGTACAA ATTCACCCGT CCGCCCACAG
701 GCGACGTGTT GCGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGGCTCCAGC TCCGTTCCGC AAGAATGGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG TAAAGTCAAC
901 ATCGATACCG ACTTGCGCCT TGCTTCCACC GGCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TCGATCCGCG CAAATATTTG AGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCCTGACC GCTACCTCGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAT CAAACCGGTT TCCTTGGAAG AAATGGCAAA
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

```

a034.pep
  1 MSRLWFFAAK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
 51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQRSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAVGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAI AVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL SKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMANRYA KGELNQIVK*

```

m034/a034 96.9% identity over a 257 aa overlap

```

          10      20      30      40      50      60
m034.pep  MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM
          || |||||:|||||
a034      MSRLWFFAAKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM
          10      20      30      40      50      60

          70      80      90      100     110     120
m034.pep  EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI
          |||||:|||||
a034      EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI
          70      80      90      100     110     120

          130     140     150     160     170     180
m034.pep  QLGFSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFHSHACGVSVEGEIGVLGNLETGDAG
          |||||:|||||
a034      QLGFSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFHSHACGVSVEGEIGVLGNLETGEAG
          130     140     150     160     170     180

          190     200     210     220     230     240
m034.pep  EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAI AVG TSHGAYKFTRPPTGDVLRID
          |||||:|||||
a034      EEDGVGAVGKLSHDQMLTSVEDAVRFVKDTGVDALAI AVG TSHGAYKFTRPPTGDVLRID
          190     200     210     220     230     240

          250
m034.pep  RIKEIHQALPNTTHIVMH
          |||||
a034      RIKEIHQALPNTTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVKRVN
          250     260     270     280     290     300

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from *N. gonorrhoeae*:

m034/g034

m034 . pep	MSCLWFFAVKNI IIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNXLQMRIM	60
g034	MSRLWFFAVKNI IIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNLEQMRIM	60
m034 . pep	EAADQVDAPVIVQASAGARKYAGAPFLRHLILA AVEVFPHI PVVMHQDHGASPDVCQRSI	120
g034	EAADQVNAPVIVQASAGARKYAGAPFLRHLILA AVEEFPHI PVVMHQDHGASPDVCQRSI	120
m034 . pep	QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVNFSSHACGVSVEGEIGVLGNLETGDAG	180
g034	QLGFSSVMDGSLLEDGKTPSSYEYNVNATRTVNFSSHACGVSVEGEIGVLGNLETGEAG	180
m034 . pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAI AVGTSHGAYKFTRPPTGDVLRID	240
g034	EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAI AVGTSHGAYKFTRPPTGDVLRID	240
m034 . pep	RIKEIHQALPNTHIVMH	257
g034	RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVRKVN	300

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 127>:

g036 . seq

```

1  ATGCTGAAGC CGTGTTTGGT ATACAGTGCC TGTGCGGCGG cgttgccCTGC
51  GCGGACTTCG AGCAGCAGGC GTTGCGTGCC TTCGGGCAGA TGTGCGTACC
101 AATATTCGAG CAGGGCGGAC GCAACGCCCC GTCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCTGCCGT CTTTTCCGC AAGGAAAACC TGTTCCGACG
251 GCGAAACAAG CGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGTAT CGAGCGCGGC CAGTGCGGCG CAGTCGGACG GTGAGGCTGG
351 GCGGATGTTT ATGTTCTGTC CTTCCGTTCC GCCTGTTCTT TGGCAGTCAG
401 GCGGATTTTG TTGCGGACGT AGAGCAGTTC GGCGTGTGCC GCGCCAGTTG
451 CCGGATAGCC GCGCGCAGG GCGAGCGCGA GAAATCGGC GCGGTCGCGC
501 ATATCGGGT TGCCTGAGAA GGGCGGACGG TTTTCCAGTG CGAACGCACT
551 QCCGATGCCG TCTGAAAAGA CGTACCCCTC GGGGAGGGCA ATGTCTGCCG
601 CCCTACCGAC TTGATAATCG CTCAAACGGC GCGGTTTCAG CGTGTCGAAC
651 CACGCATAAA ACACCTCGCC CACAGCGCG TCCGCAGCGG CGAGTATGCA
701 GCTTTCGCGC GCGGCGAGC AGGCGGCGG ATCGAGCGTG GGGATGCCGA
751 TTAAAGGCGT GTCGAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 128; ORF 036.ng>:

g036 . pep

```

1  MLKPCLVYSA CAAALPARTS SSRRCVPSGR CAYQYSSRAD ATPRRRHSGA
51  VAIRCSSDSS GRFCQTIKAA ILPSFSARKT CSDGETSADS NWRCVHADGL
101 QTVSSAASAA QSDGEAGRMF MFVPSVPPVL WQSGRFCCGR RAVRRVPRQL
151 RDSRRRGRAR ENRRRSAYRV CLRRADGFPV RTHCRCLKR RTPRGQCLP
201 PYRLDNRSNG GGSACRTHK TLRPYARPQ RVCSFAAAAA RRRHRAWGCR
251 LKACRTALPN LAPRRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 129>:

m036 . seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCGCTGC
51  ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCg.CGT CTTTTCCGC AAGGAAAACC TGTTCCGACG
251 GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGCGGC CAGCTCCTCA CAATCGGCAC AAACGGCACG

```

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTTC GCCTGTTCTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA ACACTTCGCC CATAAGAGCG TCCGTAGCGG CAAGGATGCA
701 GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>:

m036.pep

```

1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPXSFSAKRT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMP ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RI PPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251 LRGYQTALPN PELHRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 131>:

a036.seq

```

1 ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCTGCTG
51 ACGGACTTCG AGCAGCAGGC GTTGCCTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCGCGCGT CTTTTCGCGC AAGGAAAACC TGTTCGGACG
251 GCGAAAACAG TCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGCAT CGAGCGCGGC GAGTGGCGCG CAATCGGCAT AAACGGCGCG
351 GCGGATGTTT ACAGGCGCGC CCTCCGTTCC GCCTGTTCTT TGGCAGTCAA
401 GGCGATTTTG TTGCGGACGT AGAGCAGCTC GGCGTGTGCC GCAGCGACGG
451 CGGGAACACC GCCTTCAGCC GCCAGATTGA GGAAGTCGGC GGCGGTGCGC
501 ATATCGGGTT TGCTGAGAA GGGCGGACGG TTTTCCAGCG CGAACGCATT
551 GCCGATGCCG TCTGAAAAGG CGCATCCTTC CGGCAGCCGG ATGTCTGCCG
601 CCCGACCGAC CTGATAATCG CTCAAACGGC GGCAGTTCAG CGTGTGGAAC
651 CATGCATAAA ACACTTCGCC CATACTGCG TCCGAGCGG CAAGGATGCA
701 GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAAGGAGT ATCAAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

a036.pep

```

1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPPSFSAKRT CSDGETSADS NWRCVHADGL
101 QTASSAASAA QSA*TARRMF TGAPSVPPVL WQSRRFCCGR RAARRVPQRR
151 RENRLQPPD* GSRRRSAYRV CLRRADGFPA RTHCRCLKR RILPAAGCLP
201 PDRPDNRSNG GGSACRTMHK TLRPYVRPQR QGCSFAAAAA RRRHRARVRR
251 LKEYQTALPN LAPRRCRYAV P*

```

m036/a036 85.6% identity over a 270 aa overlap

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIRCSSDSS					
a036	MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSAKRTCSNWRVHADGLQTASSAASSSQSAQTARRMF					
a036	GRFCQTIKAAIPPSFSAKRTCSNWRVHADGLQTASSAASAAQSXTARRMF					
	70	80	90	100	110	120

221

	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCLRTAELLPA					
a036	TGAPSVPPVLWQSGRRFCCGRRARRVPQRRRENRLQPPDXGSRRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPNDNRNNGGSSAYRTMHKTLRPYERXPXQGCSTFAAAAA					
a036	RTHCRCLKRRILPAAGCLPPDRPDNRNNGGGSACRTMHKTLRPYVRPQRQGCSTFAAAAA					
	190	200	210	220	230	240
	250	260	270			
m036.pep	RRRHRARVRRLRGYQTALPNPELHRCRYAVRX					
a036	RRRHRARVRRLKEYQTALPNLAPRRCRYAVPX					
	250	260	270			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from *N. gonorrhoeae*:

m036/g036

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAI PWRRHSGAVAIRCSSDSS					
g036	MLKPCLVYSACAAALPARTSSSRRCVPSGRCA YQYSSRADATPRRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSARKTCS DGETSADSNWRCVHADGLQTASSAASSQSAQTARRMF					
g036	GRFCQTIKAAILPSFSARKTCS DGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCLRTAELLPA					
g036	MFVPSVPPVLWQSGRFCCGRRARRVRPQRLRDSRRRGRARENRRRSAYRVCLRRADGFVP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPNDNRNNGGSSAYRTMHKTLRPYERXPXQGCSTFAAAAA					
g036	RTHCRCLKRRTPRGQCLPPYRLDNRSNNGGGSACRTTHKTLRPYARPQRVCSTFAAAAA					
	190	200	210	220	230	240
	250	260	270			
m036.pep	RRRHRARVRRLRGYQTALPNPELHRCRYAVRX					
g036	RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 133>:

m036-1.seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTCCTGCTGC
51  ACGGACTTCG AGCAGCAGGC GTTGCCTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
251 GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGCGGC CAGCTCCTCA CAATCGGCAC AAACGGCACG

```

222

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTTC GCCTGTTCTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA AACTTCGCC CATACGAGCG TCCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>:

```

m036-1.pep
  1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51 VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMP ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*

```

m036-1/g036 76.8% identity in 228 aa overlap

```

              10      20      30      40      50      60
m036-1.pep  MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIWRRHSGAVAIIRCSSDSS
              ||||| |||||:||||| ||||| ||||| ||||| ||||| |||||
g036         MLKPCLVYSACAAALPARTSSSRRCVPSGRCAVQYSSRADATPRRRHSGAVAIIRCSSDSS
              10      20      30      40      50      60

              70      80      90      100     110     120
m036-1.pep  GRFCQTIKAAIPPSFSARKTCSDETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g036         GRFCQTIKAAILPSFSARKTCSDETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF
              70      80      90      100     110     120

              130     140     150     160     170     180
m036-1.pep  TGALSVRPVLWQSGRFCCGRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCLRTAELLPA
              : || ||||| ||||| ||| : :|| ||||| ||||| ||| : ||:
g036         MFVPSVPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPV
              130     140     150     160     170     180

              190     200     210     220     229
m036-1.pep  RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPX
              ||| ||||| | : ||||| ||||| || ||||| ||||| ||
g036         RTHCRCLKRRTPRGGQCLPPYRLDNRSNGGGSACRTHKTLRPYARPQRRVCSFAAAAA
              190     200     210     220     230     240

g036         RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX
              250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 135>:

```

g038.seq
  1 ATGACTGATT TCCGCCAAGA TTTCTCAAA TTCTCCCTCG CCCAAATGT
 51 TTGAAATTC GGCGAATTA CCACCAAAGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTC AACGACGGCG CGTCCACGCT GCAACTGGCA
151 AAATCTATG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC GCGGCGAAGG CGGCGTGTG GTGCGGCGCG CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCAATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ACCGCGATGA AAAAGGCACG GGTAATTGT CCGCGTTCA
501 GGAAGTGGA AAACAATACG GCCTGCCCCG CGCCCCCATC GCCAGCCTGA
551 ACGATTTGTT TATCTGTGTT CAAAACAACC CCGAATTCGG ACAGTTCCTC
601 GAACCCGTCC GCACCTACCG CCGGCAGTAC GCGGTAGAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:

```
g038.pep
  1  MTDFRQDFLK  FSLAQNVLFK  GEFTTKAGRR  SPYFFNAGLF  NDGASTLQLA
 51  KFYAQSIIES  GIRFDMFLGP  AYKGIILAAA  TAMMLAEKGV  NVPFAYNRKE
101  AKDRGEGGVL  VGAPLKGRVL  IIDDVISAGT  SVRESIKLIE  AEGATPAGVA
151  IALDRMEKGT  GKLSAVQEVE  KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRTYRRQY  GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 137>:

```
m038.seq
  1  ATGACCGATT  TCCGCCAAGA  TTTCTCTCAA  TTCTCCCTCG  CCCAAAATGT
 51  TTTGAAATTC  GGCGAATTTA  CCACCAAGGC  AGGACGGCGG  TCGCCCTATT
101  TCTTCAATGC  CGGCCTCTTT  AACGACGGCT  TGTCCACGCT  GCAACTGGCA
151  AAATTTTACG  CACAATCCAT  CATTGAAAGC  GGCATCCGAT  TCGATATGCT
201  GTTCGGTCCC  GCCTACAAAG  GCATTATTTT  GGCGGCGGCA  ACCGCGATGA
251  TGCTGGCGGA  AAAAGGCGTG  AACGTCCCGT  TTGCCTACAA  CCGCAAAGAA
301  GCCAAAGACC  ACGGCGAAGG  CGGCGTGTG  GTCGGCGCGC  CGCTTAAAGG
351  GCGCGTGCTG  ATTATCGACG  ACGTGATTTC  CGCCGGCACA  TCCGTACGCG
401  AATCGATCAA  ACTGATTGAA  GCGGAGGGTG  CAACCCCGCG  CGGTGTCGCC
451  ATCGCGCTCG  ATCGCATGGA  AAAAGGCACG  GGTGAATTGA  GCGCGGTTCA
501  GGAAGTGGAA  AAACAATACG  GkCTGCCCGT  CGCCCCCATC  GCCAGCTGA
551  ACGATTGTGT  TATTCTGTTG  CAAAACAACC  CCGAATTCGG  ACAGTTCCTC
601  GAACCCGTCC  GAGCCTACCG  TCGGCAGTAC  GCGGTAGAAT  AA
```

This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:

```
m038.pep
  1  MTDFRQDFLK  FSLAQNVLFK  GEFTTKAGRR  SPYFFNAGLF  NDGLSTLQLA
 51  KFYAQSIIES  GIRFDMFLGP  AYKGIILAAA  TAMMLAEKGV  NVPFAYNRKE
101  AKDHGEGGVL  VGAPLKGRVL  IIDDVISAGT  SVRESIKLIE  AEGATPAGVA
151  IALDRMEKGT  GELSAVQEVE  KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRAYRRQY  GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 139>:

```
a038.seq
  1  ATGACCGATT  TCCGCCAAGA  TTTCTCTCAA  TTCTCCCTCG  CCCAAAATGT
 51  TTTGAAATTC  GGCGAATTTA  CCACCAAAGC  CGGACGGCGG  TCGCCCTATT
101  TCTTCAATGC  CGGCCTCTTT  AACGACGGCT  TGTCCACGCT  GCAACTGGCA
151  AAATTTTACG  CACAATCCAT  CATTGAAAGC  GGCATCCGAT  TCGATATGCT
201  GTTCGGCCCC  GCCTACAAAG  GCATTATTTT  GGCGGCGGCA  ACCGCGATGA
251  TGCTGGCGGA  AAAAGGCGTG  AACGTCCCGT  TTGCCTACAA  CCGCAAAGAA
301  GCCAAAGACC  ACGGCGAAGG  CGGCGTGTG  GTCGGCGCGC  CGCTTAAAGG
351  GCGCGTGCTG  ATTATCGACG  ACGTGATTTC  CGCCGGCACA  TCCGTACGCG
401  AATCGATCAA  ACTGATTGAA  GCGGAGGGTG  CAACCCCGCG  CGGTGTCGCC
451  ATCGCGCTCG  ACCGCATGGA  AAAAGGCACG  GGTGAATTGA  GCGCGGTTCA
501  GGAAGTGGAA  AAACAATACG  GCCTGCCCGT  CGCCCCCATC  GCCAGCTGA
551  ACGATTGTGT  TATTCTGTTG  CAAAACAACC  CCGAATTCGG  ACAGTTCCTC
601  GAACCCGTCC  GAGCCTACCG  TCGGCAGTAC  GCGGTAGAAT  AA
```

This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:

```
a038.pep
  1  MTDFRQDFLK  FSLAQNVLFK  GEFTTKAGRR  SPYFFNAGLF  NDGLSTLQLA
 51  KFYAQSIIES  GIRFDMFLGP  AYKGIILAAA  TAMMLAEKGV  NVPFAYNRKE
101  AKDHGEGGVL  VGAPLKGRVL  IIDDVISAGT  SVRESIKLIE  AEGATPAGVA
151  IALDRMEKGT  GELSAVQEVE  KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRAYRRQY  GVE*
```

m038/a038 100.0% identity over a 213 aa overlap

```

              10      20      30      40      50      60
m038.pep  MTDFRQDFLKFSLSAQNVLFKGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
```

224

```

a038      |||||
          MTDFRQDFLKFSLAQNVLFKGFETTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
          10      20      30      40      50      60

          70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          |||||
a038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          70      80      90      100     110     120

          130     140     150     160     170     180
m038.pep  IIDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          |||||
a038      IIDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          130     140     150     160     170     180

          190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          |||||
a038      ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          190     200     210

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from *N. gonorrhoeae*:

m038/g038

```

m038.pep  10      20      30      40      50      60
          MTDFRQDFLKFSLAQNVLFKGFETTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
          |||||
g038      MTDFRQDFLKFSLAQNVLFKGFETTKAGRRSPYFFNAGLFNDGASTLQLAKFYAQSIIES
          10      20      30      40      50      60

          70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          |||||
g038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDRGEGGVLVGAPLKGRVL
          70      80      90      100     110     120

          130     140     150     160     170     180
m038.pep  IIDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          |||||
g038      IIDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGKLSAVQEVEKQYGLPVAPI
          130     140     150     160     170     180

          190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          |||||
g038      ASLNDLFILLQNNPEFGQFLEPVRTYRRQYGVEX
          190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 141>:

g039.seq

```

1  ATGCCGTC CG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACAC
51 CGAGAAAAACA TCATGCCCGC CTGTTTCTGT CCGCACTGCA AAACCCGCCT
101 CTGGGTCAAA GAAAcccagC TCAAcgtCgC ccaagGCTTC GTCGTCTgcc
151 aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggc aaGcacGAAA
201 gaacctatat tcaacgattg gcccgaaagt gtttcgggat gTcaaaCTCG

```

225

```

251 TCaccgcgaT cggcacgcac gccattagca aGAaacagat gtccccgcgac
301 gaaatCgccg atatacctcaa cggcggtaca acCCTGCACG ATACGCCCGCC
351 CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGCaggTT TCCGTACCGC
401 CCGCCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
451 ATCGTCCTCA TTATGCAGCT TTCCTACCTC TTCATCCTAT GA

```

This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:

g039.pep

```

1  MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
51  KNAKGLKPK TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKQMSRD
101 EIADILNGGT TLHDTFPATA AAAPAAAPQV SVPPARQEGT NWTIATLFAL
151 IVLIMQLSYL FIL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 143>:

m039.seq

```

1  ATGCCGTCCG AACC GCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51  CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCC ACTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCCAAC TCAATGTTCG CGnnnnnnnnn nnnnnnnnnn
151 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnCCC GAGGCTGTTT
251 CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
301 CAGATTTCCC GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
351 GCCCGATATT CCGCCGCAA CCGCCGCCAC CCCTGCTGCC GCACCGCAGG
401 TTACCGTACC GCCCGCGCG CCCGCCGTC AGGATGGGT CAACTGGACG
451 ATTGCAACCC TGTTCGCCCT TATCGTCCTC ATTATGCAGC TTTCTACCT
501 CGTCATCCTA TGA

```

This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:

m039.pep

```

1  MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPPXXXXXX
51  XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX EAVSDVKLVH RIGTRAIGKK
101 QISRDEIAGI LNGGTTQPD I PPATAATPAA APQVTVPAA PARQDGFNWT
151 IATLFALIVL IMQLSYLVIL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 145>:

a039.seq

```

1  ATGCCGTCTG AACC GCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51  CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCC ACTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCCAAC TCAATGTTCG CCAAGGCTTC GTCGTCTGCC
151 AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCACGAAA
201 GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTTCGGAT GTCAAACTCG
251 TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
301 GAAATCGCCG GCATCCTCAA CGGCGGCACA ACCCAGCCCG ATATTCCGCC
351 CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
401 CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCTGTTT
451 GCCCTTATCG TCCTCATTAT GCAGCTTTC TACCTCGTCA TCCTATGA

```

This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:

a039.pep

```

1  MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
51  KNAKECLKPK TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISR
101 EIAGILNGGT TQPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLF
151 ALIVLIMQLS YLIVL*

```

m039/a039 79.4% identity over a 170 aa overlap

```

              10      20      30      40      50      60
m039.pep      MPSEPPYASDGIKPDTHEEIPCPPVSAPTA KPVSGSKKPN SMSPPXXXXXXXXXXXXXXXXX
               |||
a039           MPSEPPYASDGIKPDTHEEIPCPPVSAPTA KPVSGSKKPN SMSPKASSSAKNAKECLKPK

```


Homology with a predicted ORF from *N. gonorrhoeae*

m039/g039

```

                                10      20      30      40      50      60
m039.pep  MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVGSGSKKPNSMSPXXXXXXXXXXXXXXXXXX
          ||||| ||||| : ||||: |||: ||||: |||
g039      MPSEPPAASDGIKPTHTEKTSCPPVSVRTAKPASGSKKPSSTSPKASSSAKNAKGCLKPK
          10      20      30      40      50      60

                                70      80      90      100     110     120
m039.pep  XXXXXXXXXXXXXXXXXXXXPEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPD
          :      :      : ||||| |||||: ||: ||: ||||| ||||| |||||
g039      TIWQARKNLYSTIG-----PKLFRDVKLVHRIGTHAISKQMSRDEIADILNGGTTLHDT
          70      80      90      100     110

                                130     140     150     160     170
m039.pep  PPATAAT-PAAAPQVTVPAAAPARQDGFNWTIATLFLALIVLIMQLSYLVILX
          |||||: |||||: ||| |||: ||: ||: ||||| ||||| |||||
g039      PPATAAAPAAAPQVSVPPA---RQEGLNWTIATLFLALIVLIMQLSYLFILX
          120     130     140     150     160

```

q040.seq

1	ATGAACGCGC	CCGACAGCTT	TGTCGCCAC	TTCCGCGAAG	CCGCCCCCTA
51	CATCCGCCAA	ATGCGCGGCA	CGACACTGGT	CGCCGGCATA	GAaggCCGCC
101	TGCTCGAAGG	CGGCACCTTA	AATAAGCTCG	CCGCCGACAT	CGGGCTGTTG
151	TCGCAACTGG	GCATCCGACT	CGTCCCTCAT	CACGGCGCGT	ACCACCTTCT
201	CGAacgCCTC	GCCGCCGCGC	AAGgcccGCAC	CGCGCATTAT	TGCCGgggtt
251	tGCGCGTTAC	CGACGaAAcc	tcGctcgGAC	AGGCGCAGCA	GtttGCCGGC
301	AccgTCCGCA	GCCGTTTTGA	agcCGCATTG	tgcggcagCG	tttcaggatt
351	cgcgCGCGCG	CCTTCCGTCC	CGCTCGTAtc	gggcaacttc	ctgacCGCCC
401	GTCCgattggg	cgtgattgac	ggaACCGata	tgaataatcg	gggggttatc
451	gcgaaaaccg	ACACGCCCGC	CCTCCGTTTC	CAGTCCGACG	CGGGCAATAT
501	CGTCTGGATG	CCGCCGCTCG	GGCATTCCTA	CGGCGGCAAA	ACCTTCAATC
551	TCGATATGGT	GCAGGCCGCC	GCTTCCGTCG	CGGTCTCGCT	TCAGGCCGAA
601	AAACTCGTTT	ACCTGACCTT	TTCAGACGGC	ATTTCCCGCC	CCGACGGCAC
651	GCTCGCCGAA	ACCCTCTCGG	CACAGGAAGC	GCAATCGCTG	CGGGAACACG
701	CCGCCACGCA	AACCCAGCA	CTGATTTCGT	CGCCGGTTGC	CGCGCTCGAA
751	GGCGGCGTGC	ATCGCGTCCA	AATCCTCAAC	GGGGCCGCCG	ACGGCAGCCT
801	GCTGCAAGAA	CTCTTCACCC	GCAACGGCAT	CGGCACGTCC	ATTGCCAAAG
851	AAGCCTTCGT	CTCCATCCGG	GAGGCGCACA	GCGGCGACAT	CCCGCACATC
901	GCCGCCCTCA	TCCGCCCGCT	GGAAGAACAG	GGCGTCCAT	TGCACCGCAG
951	CCGCGAATAC	CTCGAAAACC	ACATTTCGGA	ATTTTCCATC	CTCGAACACG

```

1001 ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051 TCGGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGg
1101 ctACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151 GCATAAGCAG GCTGTTGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201 GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCG AAACGCGGCG
1251 CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301 TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>:

g040.pep

```

1 MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DGRLLGGTLL NKLAADIGLL
51 SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGTALAE TLSAQEAQSL AEHAASETRR LISSAVAALE
251 GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GVLHRSREY LENHISEFSI LEHDGDLYGC AALKTFAEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNGRNPHIL VRRLLHR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 149>:

m040.seq

```

1 ATGAGCGCGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGTCCCCTA
51 CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACGCC
101 TGTCGAAGG TGATACCTTA AACAGCTCG CCGCCGACAT CGGGCTGTTG
151 TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TCGCGGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
301 ACCGTCCGCA GCCGTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
351 CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAGG TGTGATTGAC GGAACGATA TGGAATACGC GGGCGTTATC
451 CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
501 CGTCTGGCTG CCGCCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCTATC
551 TCGATATGCT TCAAACCGCC GCCTCCGCCG CCGTCTCGCT TCAGGCCGAA
601 AAATCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651 GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACAG
701 CCGCGGGGCA AACGCGACGG CTGATTTCGT CCGCCGAATC CTTCACCCGC
751 AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
801 rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
851 AAGAACAGGG CATCTGCTG CACCGCAs.c GCGAATACCT CGAAAACCA
901 ATTTCCGAAT TTTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
951 CGCCCTGAAA ACCTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001 CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
1051 CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTGCGACT
1101 GTCCACAAAT ACCGGCGAAT GGTTTGCCGA ACGCGGCTTT CAGACGGCAT
1151 CGGAAGACGA GTTGCCCGAA ACGCGGCGCA AAGACTACCG CAGCAACGGA
1201 CGGAATCGC ATATTCTGGT ACGTCGCCTG CACCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

m040.pep

```

1 MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLGDTLL NKLAADIGLL
51 SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201 KLVYLTLSDG ISRPDGTALAE TLSAQEAQSL AEHAGGQTRR LISSAELFTR
251 NGIGTSIAKE AFVSIRQAHX XDIPHIALI RPLEEQGILL HRXREYLENH
301 ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351 HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401 RNSHILVRRL HR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 151>:

a040.seq

a040.pcp

1	MIVPDLFVAH	FREAAPYIRQ	MRGKTLVAGI	DDRLLGDTL	NKFAADIGLL
51	<u>SQLGIRLVLI</u>	HGARHFLDRH	AAAQGRTPHY	CRGLRVTDET	SLEQAQQFAG
101	<u>TVRSRFEAAL</u>	CGSVSGFARA	PSPVLVSGNF	LTARPIQIDA	GTPMEYAGVI
151	RKTDTAALRF	QLDAGNIVWL	PPLGHSYSGK	TFHLDMLQTA	ASVAVSLQAE
201	KLVLTLSDG	ISRPDGT LAV	TLSAQEAQSL	AEHAGGETRR	LISSAVAALAE
251	GGVHRVQILN	GAADGSLLQE	LFTRNGIGTS	IAKEAFVSIR	QAHSGDIPHI
301	AALIRPLEEQ	GILLHRSREY	LENHISSEFSI	LEHDGNLYGC	AALKTFAEAD
351	CGEIACLAS	PQAQDGGGYE	RLLAHIIDKA	RGISIRLFA	LSTNTGEWFA
401	ERGFOTASED	ELPETRRKDY	RSNGRNSHIL	VRRLHR*	

m040/a040 91.5% identity in 436 aa overlap

	10	20	30	40	50	60
m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI					
a040	MIVPDLFVAHFREAAPYIRQMRGKTLVAGIDDRLLEGDTLNKFAADIGLLSQLGIRLVLI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m040.pep	HGARHFLLDRHAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA					
a040	HGARHFLLDRHAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPLPLGHSYS GK					
a040	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPLPLGHSYS GK					
	130	140	150	160	170	180
	190	200	210	220	230	240
m040.pep	TFYLDMLQTAASAASVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR					
a040	TFHLDMLQTAASVAVSLQAEKLVYLTLSDGISRPDGTLAVTLAQEAQSLAEHAGGETRR					

229

	190	200	210	220	230	240
m040 . pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI					
a040	LISSAVAALGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI					
	250	260	270	280	290	300
m040 . pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS					
a040	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS					
	310	320	330	340	350	360
m040 . pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY					
a040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY					
	370	380	390	400	410	420
m040 . pep	RSNGRNSHILVRRLHRX					
a040	RSNGRNSHILVRRLHRX					
	430					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from *N. gonorrhoeae*:

m040/g040

m040 . pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI	60
g040	MNAPDSFVAHFREAAPYIRQMRGTTLVAGIDGRLLGGTTLNKLAADIGLLSQLGIRLVLI	60
m040 . pep	HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA	120
g040	HGAYHFLDRLAAAQGRTPHYCRGLRVTDETSLGQAQQFAGTVRSRFEAALCGSVSGFARA	120
m040 . pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPLGHSYSGK	180
g040	PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYGGK	180
m040 . pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSGDGISRDPDGLAETLSAQEAQSLAEHAGGQTRR	240
g040	TFNLDMVQAAASVAVSLQAEKLVYLTLSGDGISRDPDGLAETLSAQEAQSLAEHAASETRR	240
m040 . pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI	276
g040	LISSAVAALGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI	300
m040 . pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS	336
g040	AALIRPLEEQGVLLHRXREYLENHISEFSILEHDGDLGCAALKTFAEADCGEIACLAVS	360
m040 . pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY	396
g040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY	420
m040 . pep	RSNGRNSHILVRRLHRX	413
g040	RSNGRNPILVRRLHRX	437

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 153>:

```
g041.seq
  1 ATGAGTTTCGC CCAAACACAT CGGCTTGCAG GCGGCGCAGCA ACGGCGGCCT
 51 GATTACCGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGTGCGCTGG
101 TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
151 GGTTCAAGTT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
201 CAAACGCCCG CTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCACTC ATTACCACCA GCCTCAGCGA CGACCGCGTC
301 CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCAACGCGA ATCCGCCGAC AAACCTGCCT GCGTGTGCT GTTTTGTAAA
451 GAATTTTGG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

```
g041.pep
  1 MSSPKHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
 51 GSSWTDEYGN PQKYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQRESAD KLACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 155>:

```
m041.seq
  1 ATCAGTTTCGC CCGAACACAT CGGCTTGCAG GCGGCGCAGCA ACGGCGGACT
 51 GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCCG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCCGC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCAACGCGA ATCCGCCGAC GAACCTGCCT GCGTCTTGCT GTTTTGTAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

```
m041.pep
  1 ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
 51 GSSWTDEYGN PQKYEVCRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 157>:

```
a041.seq
  1 ATCAGTTTCGC CCGAACACAT CGGCTTGCAG GCGGCGCAGCA ACGGCGGACT
 51 GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATA GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCCG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CGCAGCGCA AGCCGCCGAC GAACCTGCCT GCGTGTGCT GTTTTGTAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

```
a041.pep
  1 ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
 51 GSSWTDEYGN PQKYEVCRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQREAA ELACVLLFLK
```

151 EFLG*

m041/a041 98.7% identity over a 154 aa overlap

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
a041	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSAQSW					
a041	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSAPQSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
a041	LYSPDGGGHTGNGTQREAADELACVLLFLKEFLGX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 041 shows 96.8% identity over a 154 aa overlap with a predicted ORF (ORF 041.ng) from *N. gonorrhoeae*:

m041/g041

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	:					
g041	MSSPKHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSAQSW					
g041	PQKYEACKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSAPQSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
g041	LYSPDGGGHTGNGTQRESADKLACVLLFLKEFLGX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 159>:

g041-1.seq

1	ATGAAATCCT	ACCCCGACCC	CTACCGCCAT	TTTGAAAACC	TCGATTCCGC
51	CGAAACGCAA	AACTTCGCTG	CTGAAGCGAA	TGCCGAAACG	CGCGCGCGTT
101	TTTTAAACAA	CGACAAGGCG	CGCGCACTTT	CAGACGGCAT	TTTGAATCAA
151	ATGCAGGACA	CGCGGCAGAT	TCCGTTTTGT	CAGGAACACC	GCGCGCGGAT
201	GTACCATTTT	CATCAGAATG	CGGAATATCC	GAAGGGCGTG	TACCGCATGT
251	GACCGGCGGC	GACCTACCGT	TCCGGCTATC	CCGAGTGGA	AATCCTGTTT
301	TCGGTGGCGG	ATTTTCGATGA	GTTGCTCGGC	GACGATGTGT	ATTTGGGCGG
351	CGTGTCGCAC	TTGGTGGAGC	AGCCCAACCG	CGCGCTGCTG	ACTTTGAACA
401	AATCGGGCGG	CGATACGGCG	TATACGCTGG	AAGTGGATTT	GGAAGCAGGG
451	GAATTGGTAG	AGGGCGGTTT	TCACTTTCCG	GCAGGCAAAA	ACCATGTGTC
501	GTGGCGCGAT	GAAAACAGCG	TGTGGGTGTG	TCCGGCTTGG	GACGAACGCC
551	AGTTGACCGA	ATCGGGCTAT	CCGCGCGAAG	TGTGGCTGGT	GGAACGCGGC

```

601 AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
651 GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCTG
751 TCCGAAGGCG GGGCGAAACC GTTGAACCTG CCTAATGATT GCGATGTGGT
801 CGGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAAACTGAAT
901 CCGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
1001 TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051 TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTC
1151 CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAACGTACC
1201 GTCATGCGCC TCCAGCCGCA GCAGTTTGTT TCAGACGGCA TCGAAGTGCG
1251 GCAGTTTGG GCGGTGTCTG CCGACGGCGA ACGCATTCCT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1351 GGTTTCGGCA TTCCTGAATT GCCGCATTAT CTGGGCAGCG TCGGCAATA
1401 TTGCTGGAA GAGGGCAATG CCTTTGTATT GGCAAACATC CGCGGCGGCG
1451 GAGAATTCCG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1501 AAAAGCGTTG ATGATTGTGT GGCAGTCGTG CGTGATTGTG CCGAACGCGG
1551 CATGAGTTCG CCCAAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGCC
1601 TGATTACCGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT CGGTGCGCTG
1651 GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701 CGGTTCAAGT TGGACGGACG AATACGGCAA TCCGCAGAAA TACGAAGCCT
1751 GCAAACGCCG GCTGGGCGAA TTGTGCGCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCGT
1851 CCATCCGCCC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CAAACTCGCC TGCGTGTTCG TGTTTTTGAA
2001 AGAATTTTTG GGATAA

```

This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

g041-1.pep

```

1 MKSYDPYRHH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
51 MQDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLNKSGGDTA YTLEVDLEAG
151 ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 SEGGAAPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSYF SGALVAVKLN
301 RGELGAAQLL FAPDETQALE SVETTKRFVY ASLLENVQGR LKAWRFADSK
351 WQEAELPHLP SGALEMTDQP WGGDVVYLAA SDFTTPTLTF ALDLNVMELT
401 VMRLQPQQFV SDGIEVRQFW AVSSDGERIP YFHVGNKNAAP DTPTLVYAYG
451 FGFIPELPHY LGSVGGKYWL EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLLAVV RDLSEGRMSS PKHIGLQGGG NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQRESADKLA CVLLFLKEFL G*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 161>:

m041-1.seq

```

1 ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51 CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAGAAAA CGACAAGGCG CGCGCGCTTT CAGACGGCAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAGGACG CGGAGTATCC GAAGGGCGTG TACCGCGTGT
251 GTACCGCGGC GACGTATCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTTCGACGA ATTGCTTGGC GACGATGTGT ATTTGGGCGG
351 CTGTGCGCAC TTGGTGGAAC AGCCCAACCG CGCGTTGTTA ACGTGGAGCA
401 AATTGGGCAG CGATACGGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTCG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG AACGAACGCC
551 AGTTGACCCA ATCGGGCTAT CCGCGCGAAG TATGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT CCGTGTGTAT CAAATCGGCG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ATCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AAACCTATTT GCGGGTCTCA

```

```

751 GCCGAAGGCG AGGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTGGT
801 CGGCTATCTG GCGGGGCATC TTTTGCTGAC GCTGCGCAAG GACTGGAACC
851 GCGCGAACCA AAGCTATCCG AGCGGCGCGC TGGTGGCGGT GAAGCTGAAT
901 CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCGAGCCTGT
1001 TGGAGAACGT ACAAGGCCGT CTGAAAGCAT GGCGGTTTGC CGACGGCAAA
1051 TGGCAGGAAG TCGAATTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCT TGGGGCGGCG ACGTGGTTTA CCTTGCCGCC AGCGATTTCA
1151 CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAGTACC
1201 GTCATGCGCC GCCAGCCGCA GCAGTTGAT TCAGACGCA TTAACGTGCA
1251 GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC
1351 GGTTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA
1401 TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCGAACATC CGCGGCGGCG
1451 GCGAGTTCCG CCCGCGCTGG CATCAGCGCG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTATT GGCAGTCGTG CCGGATTTGT CCGAACGCGG
1551 TATCAGTTTC CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGCGCGAC
1601 TGATTACTGC CGCCGCTTC GTGCGCGAAC CGCAAAGCAT CGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCCG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCCG
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CGAACTCGCC TCGCTCTTGC TGTTTTTGAA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 162; ORF 041-1>:

m041-1.pep

```

1  MKSYDPYRH FENLDSAETQ NFAAEANAET RARFLENDKA RALSDGILAQ
51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKLGSDTA YTLEVDLEAG
151 ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW NERQLTQSGY PREVWLVERG
201 KSFEESLPVY QIGEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLRVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLLTLRK DWNRRANQSY SGALVAVKLN
301 RGELGAAQLL FAPDETQALE SVETTKRFV ASLLENVQGR LKAWRFADGK
351 WQEVLPRLP SGALEMTDQF WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGNKNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLAVV RDLSESGISS PEHIGLQGGG NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSASQSWLYS PDGGGHTGNG
651 TQRESADELA CVLLFLKEFL G*

```

m041-1/g041-1 94.6% identity in 671 aa overlap

	10	20	30	40	50	60
m041-1.pep	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLENDKARALSDG	ILAQ	LQDTRQIPFC	
g041-1	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLNNDKARALSDG	ILNQMQDTRQIPFC		
	10	20	30	40	50	60
m041-1.pep	QEHRARMYHFHQDAEYPKGV	YRVCTAATYRSGYPEWKILF	SVADFDELLGDDVYLGGVSH			
g041-1	QEHRARMYHFHQNAEYPKGV	YRMCTAATYRSGYPEWKILF	SVADFDELLGDDVYLGGVSH			
	70	80	90	100	110	120
m041-1.pep	LVEQPNRALLTLSKLGSDTA	YTLEVDLEAGELVEGGFHFP	AGKNHVSWRDENS	VWVCPAW		
g041-1	LVEQPNRALLTLNKS	GGDTAYTLEVDLEAGELVEGGFHFP	AGKNHVSWRDENS	VWVCPAW		
	130	140	150	160	170	180
m041-1.pep	NERQLTQSGYPREVWLVERG	KSFEESLPVYQIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG		
g041-1	NERQLTQSGYPREVWLVERG	KSFEESLPVYQIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG		
	190	200	210	220	230	240
m041-1.pep						
g041-1						

234

```
g041-1      DERQLTESGYPREVWLVERGKSFEESLPAYQIDKGAMMVNAWRYLDPQGSPIDLIEASDG
              190      200      210      220      230      240
m041-1.pep   250      260      270      280      290      300
              FYTKTYLRVSAEAGEAKPLNLPNDCDVVGYLAGHLLLTIRKDWNRANQSYPSGALVAVKLN
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      FYTKTYLQVSSEGGAKPLNLPNDCDVVGYLAGHLLLTIRKDWNRANQSYPSGALVAVKLN
              250      260      270      280      290      300
m041-1.pep   310      320      330      340      350      360
              RGELGAAQLLFAPDETOALESVETTKRFVVASLLENVQGRKAWRFADGKWQEVELPRLP
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      RGELGAAQLLFAPDETOALESVETTKRFVVASLLENVQGRKAWRFADSKWQEAELPHLP
              310      320      330      340      350      360
m041-1.pep   370      380      390      400      410      420
              SGALEMTDQPWGGDVVYLAASDFTTPTLTLFALDLNVMELTMRRQPQQFSDGINVQQFW
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      SGALEMTDQPWGGDVVYLAASDFTTPTLTLFALDLNVMELTMRLQPQQFVSDGIEVRQFW
              370      380      390      400      410      420
m041-1.pep   430      440      450      460      470      480
              TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI
              ::|:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      AVSSDGERIPYFHVGNKNAAPDTPTLVYAYGGFGIPELPHYLGSVKGKYWLEEGNAFVLANI
              430      440      450      460      470      480
m041-1.pep   490      500      510      520      530      540
              RGGGEFGPRWHQAAQGISKHKSVDLLAVVRDLSESGISSPEHIGLQGGSNGLITAAAF
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      RGGGEFGPRWHQAAQGISKHKSVDLLAVVRDLSESGMSSPKHIGLQGGSNGLITAAAF
              490      500      510      520      530      540
m041-1.pep   550      560      570      580      590      600
              VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCRRRLGELSPYHNLSDG
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEACKRRRLGELSPYHNLSDG
              550      560      570      580      590      600
m041-1.pep   610      620      630      640      650      660
              IDYPALITTSLSDDRVHPAHALKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADELA
              |||||:|:| |||||:|:| |||||:|:| |||||:|:| |||||:|:|
g041-1      IDYPALITTSLSDDRVHPAHALKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADKLA
              610      620      630      640      650      660
m041-1.pep   670
              CVLLFLKEFLGX
              |||||:|:|
g041-1      CVLLFLKEFLGX
              670
```

m041-1/P55577

sp|P55577|Y4NA_RHISN PROBABLE PEPTIDASE Y4NA >gi|2182536 (AE000086) Y4nA [Rhizobium sp.
NGR234] Length = 726

Score = 370 bits (940), Expect = e-101

Identities = 217/682 (31%), Positives = 331/682 (47%), Gaps = 22/682 (3%)

```
Query: 2    KSYPDPYRHFNLD SAETQNF AAEANAETRARFL ENDKARALSDG ILAQLQDTRQIPFCQ 61
           K  DP + +D + + N T + ++ + L LQ T +I
Sbjct: 42   KDASDPRAYLNEIDGDKAMTWVEAHNLSTVDKLSKDP RYSEYQADALTILQATDRIASPS 101

Query: 62   EHRARMY-HFHQDAEYPKG VYRVCTAATYRS GYPEWKILFSVADFDELLGDDVYLGGVSH 120
           R  M +F QD + +G++R T +YRSG P+W+ + V + G G
Sbjct: 102  FARDGMIDNFWQDGHVQGLWRRTTWESYRSGNPQWR TILDVDALSKAEGKTWVFEGGDC 161

Query: 121  LVEQPNRALLTSLKLSGSDTAYTLEVDLEAGELVEGGFHFPA GKNHVSWRDENS VWVCPAW 180
           L  N L+ LS G D E D+ GE V+ GF P GK V+W DEN+++V W
```

Sbjct: 162 LPPTSNLCLIRLSDDGGKDADVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIIYVTREW 221

Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ-----IGEDGMM--VNAWRYLDPQGSPI 232
 ++T SGY +V+RG+S ++++ +++ E G++ ++ +D +

Sbjct: 222 TPGEVTSSGYAYVTKVVKRGQSLDQAVEIFRGQKKDVS AERGLRDIDGKYVMDTSYRGL 281

Query: 233 DLIEASDGFYTKTYLRVSAEGEAKPLNLPND CDVVGYLAGHLLTLRKDWN RANQS-YPS 291
 D FY + + L LP GY G + L+ DW A + + +

Sbjct: 282 DFFNTELA FY PNGH----PDTRKVV LPLPTTAVFSGY YKGQAIYWLKSDWTS AKGTVFHN 337

Query: 292 GALVAVKLN R GELGAAQL----LFAPDETQALESVETTKRFVVASLLENVQGR LKAWRFA 347
 GA++A L A++ LF P+E Q++ TK +V S+L NV ++++ F

Sbjct: 338 GAI IAFDLKAALADPARVEPLVLFMPNEHQSVAGTTQTKNRLVLSILSNVTSEVRSFDFG 397

Query: 348 DGKQWEVELPRLPSGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNMELTMRRQPQ 407
 G W +L + L +T D +++ + F P TLF D ++ + P

Sbjct: 398 KGGWSSFKLALPENSTLSLTSSDDES DQLFVVFSEGFLEPSTLFCADAATGQVEKITSTPA 457

Query: 408 QFDS DGINVQQFWTTSADGERIPYFHV GKNAA P---DMPTLVYAYGGFGIPELPHYLGSI 464
 +FD+ G+ QQFW TS DG ++PYF V + PT++YAYGGF IP P Y +

Sbjct: 458 RFDAGGLQAQQFWATSKDGT KVPYFLVARKDV KLDGNTPTILYAYGGFQIPMQPSYS AVL 517

Query: 465 GKYWLEEGNAFVLANIRGGGEFGPRWHQAAQGISKHKSVDLLAVVRDLSE RGISSPEHI 524
 GK WLE+G A+ LANIRGGGEFGP+WH A ++ + DD AV +DL + ++S H+

Sbjct: 518 GKLWLEKGGAYALANIRGGGEFGPKWH DAGLKTNRQRVYDDFQAVAQDLIAKKVTSTPHL 577

Query: 525 GLQGGSGNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEV C 584
 G+ GGSNGGL+ ++ P A+V +VPL DM+ + +SAG+SW EYG+P V

Sbjct: 578 GIMGGSNGGLLMGVQMIQRPD LWN AVVIQVPLLD MVNFT RMSAGASWQA EYGS PDD-PVE 636

Query: 585 KRRLGELSPYHNLS D GIDYPPALITTSLSDDR VHPAHALKFYAKLRETS AQSWLYSPDGG 644
 L +SPYHN+ G+ YP TS DDRV P HA K A + + Y G

Sbjct: 637 GAFLRSISPYHNVKAGVAYPEPFETSTKDDR VGPVHARKMAALFEDMGLPFY Y YENIEG 696

Query: 645 GHTGNGTQRESADELACVLLFL 666
 GH +E A A +++

Sbjct: 697 GHAAAA NLQE HARRYALEYIYM 718

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 163>:

a041-1.seq

```

1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAAACAA CGACAAGGCA CGCGCATTTGT CTGACGGCAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAAAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT
251 GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTGACGTA ATTGCTCGGT GACGATGTAT ATCTAGGCGG
351 CGTGTCGCAC CTGGTGGAAC AGCCCAACCG CGCGTTGTTA ACACTGAGCA
401 AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551 AGTTGACCGA ATCGGGCTAT CCGCGCGAGG TGTGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT GCCGGTGTAC CAAATTGCTG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA
751 GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT
801 CGGCTATCTG GCCGGACATC TTTTGCTGAC CTTGCGTAAA GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAAC
901 CGCGGCGAAT TGGGCGCGGC GCAGCTTTTG TTTGCGCCCA ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAG CGACCAAGCG TTTTGTCTGT GCGAGCCTGC
1001 TGGA AACGT ACAGGGTCGT CTGAAAGCGT GGCGTTTTAC TGATGGCAAA
1051 TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCGCC AGCGATTPCA
1151 CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACGTACC
1201 CCGATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
1251 GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC
1351 GGTTTTCGGA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA
1401 TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCGAACATC CGCGGCGGCG

```

```

1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTTATT GGCAGTCGTG AGCGATTTGT CCGAACGCGG
1551 TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT AGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCCG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACGCAGCGCG AAGCCGCCGA CGAACTCGCC TCGTGTTGC TGTTTTTGAA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>:

a041-1.pep

```

1  MKSYDPYRHFENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKSGGDTA YTLEVDLEAG
151 ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPVY QIAEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLTLRK DWHRANQSYF SGALVAVKLN
301 RGELGAAQLL FAPNETQALE SVETTKRFVV ASLLENVQGR LKAWRFTDGK
351 WQETELPRLP SGALEMTDQP WGGDVVYLA SFTTPTLTLF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGNKNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLLAVV SDSLSEGISS PEHIGLQGS NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQREAADELA CVLLFLKEFL G*

```

a041-1/m041-1 .97.9% identity in 671 aa overlap

	10	20	30	40	50	60
a041-1.pep	MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLNNDKARALSDGILAQDTRQIPFC					
m041-1	MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQDTRQIPFC					
	10	20	30	40	50	60
	70	80	90	100	110	120
a041-1.pep	QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH					
m041-1	QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH					
	70	80	90	100	110	120
	130	140	150	160	170	180
a041-1.pep	LVEQPNRALLTLSKSGGDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW					
m041-1	LVEQPNRALLTLSKLSGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW					
	130	140	150	160	170	180
	190	200	210	220	230	240
a041-1.pep	DERQLTESGYPREVWLVERGKSFEESLPVYQIAEDGMMVNAWRYLDPQGS PIDLIEASDG					
m041-1	NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVNAWRYLDPQGS PIDLIEASDG					
	190	200	210	220	230	240
	250	260	270	280	290	300
a041-1.pep	FYTKTYLQVSAEAGEAKPLNPNDCDVVGYLAGHLLTLRKDWHRANQSYPSGALVAVKLN					
m041-1	FYTKTYLRVSAEAGEAKPLNPNDCDVVGYLAGHLLTLRKDNWRANQSYPSGALVAVKLN					
	250	260	270	280	290	300
	310	320	330	340	350	360
a041-1.pep	RGELGAAQLLFAPNETQALESVETTKRFVVASLLENVQGR LKAWRFTDGKWQETELPRLP					
m041-1	RGELGAAQLLFAPDETQALESVETTKRFVVASLLENVQGR LKAWRFADGKWQEVLPRLP					
	310	320	330	340	350	360

a041-1.pep	370	380	390	400	410	420
	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMRRQPQQFDSGGINVQQFW					
m041-1	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMRRQPQQFDSGGINVQQFW					
a041-1.pep	430	440	450	460	470	480
	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI					
m041-1	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI					
a041-1.pep	490	500	510	520	530	540
	RGGGEFGPRWHQAAQGISKHKSVDLLAVVSDLSESGISSPEHIGLQGGSNGLITAAAF					
m041-1	RGGGEFGPRWHQAAQGISKHKSVDLLAVVRDLSESGISSPEHIGLQGGSNGLITAAAF					
a041-1.pep	550	560	570	580	590	600
	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
m041-1	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
a041-1.pep	610	620	630	640	650	660
	IDYPPALITTSLSDDRVRHPAHALKFYAKLRETSPQSWLYSPDGGGHTGNGTQREAADELA					
m041-1	IDYPPALITTSLSDDRVRHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADELA					
a041-1.pep	670					
	CVLLFLKEFLGX					
m041-1	CVLLFLKEFLGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 165>:

g042.seq

```

1  ATGACGATGA TTTGCTTGCG CTTCCAagcG TTCGTGCCGC ATACCAGCGC
51  GTTATCCAAC ACTTCCACGG CAGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TGCGGTCGAT GATGAAAATC CAGCCGGGGT TTTTCTCTTT GATGTATTTC
151 AAGGAAACGG GCTGCCCGTG CCCTTCGTTG CGTAAAGATT CGTCCACGGG
201 CGGCAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GATTGCGTGC
251 CGAAGGCGGA CACCTTGTG CCTGTAACG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG cggCTTCGCG CTTTTGGGCG AACAGCGCGT CAATCTGCGC
351 ATTCATATCC GCCACGCGCG CTTCTTACC GAAAATCCGC GACAGGTCT
401 CCATCTGCTT CTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAAA
451 TCTATGgtgG TCGCGTTTTT CGCCAACGT TCATACGCTT CCGCACCCGG
501 CCCGCCGTA ATGACAAACT GCGGATTGTG GCGGTGCAGG GATTGCAAT
551 CGGGCTCAAA CAGCGTCCCC ACCGTTGCCG CTTGTGCAAA TGCAGGCTGC
601 AAATAG

```

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>:

g042.pep

```

1  MTMICLRFOA FVPHTSALS N TSTAAGPSCP MAAVRSMMKI QPGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVT DSTSPRP
101 LPLAASREWA NSASICAFNS ATRASLPKIR DRVSICFSPL VRILPLSTVK
151 SMVVAFFANC SYASAPGPPV MTNCGLWRCR DSQSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 167>:

m042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCGGCCy TTCyTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CTTGTCAAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:

m042.pep

```

1 MTMICLRFOA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVT DSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 169>:

a042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAACGTG TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CTTGTCAAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:

a042.pep

```

1 MTMICLRFOA FVPRTSALSX TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVT DSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
201 K*

```

m042/a042 99.0% identity over a 201 aa overlap

	10	20	30	40	50	60
m042.pep	MTMICLRFOAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
a042	MTMICLRFOAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
m042.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
a042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
m042.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
a042	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
	130	140	150	160	170	180
m042.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
a042	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					

239

```

a042      |||||
          AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR
          130      140      150      160      170      180

          190      200
m042.pep  ASXSGSNSVPTVAALSNAGCKX
          |||||
a042      ASXSGSNSVPTVAALSNAGCKX
          190      200

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 042 shows 93.0% identity over a 201 aa overlap with a predicted ORF (ORF 042.ng) from *N. gonorrhoeae*:

m042/g042

```

          10      20      30      40      50      60
m042.pep  MTMICLRFAQFVPRTSALSXTSTAAGXSCPMMAVRSMMKIQSGFFSLMYSKETGCPCPSL
          |||||
g042      MTMICLRFAQFVPHTSALSNSTSTAAGPSCPMMAVRSMMKIQPGFFSLMYSKETGCPCPSL
          10      20      30      40      50      60

          70      80      90      100     110     120
m042.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDTSPRPLPLAASRVWANSASICAFNS
          |||||
g042      RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDTSPRPLPLAASRFWANSASICAFNS
          70      80      90      100     110     120

          130     140     150     160     170     180
m042.pep  AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR
          |:||||| :||||| :||||| :||||| :||||| :|||
g042      ATRASLPKIRDRVSICFSPLVRILPLSTVKSMVVAFFANCSYASAPGPPVMTNCGLWRCR
          130     140     150     160     170     180

          190     200
m042.pep  ASXSGSNSVPTVAALSNAGCKX
          | |||||
g042      DSQSGSNSVPTVAALSNAGCKX
          190     200

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 171>:

m042-1.seq

```

1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACC GC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA A

```

This corresponds to the amino acid sequence <SEQ ID 172; ORF 042-1>:

m042-1.pep

```

1  MTMICLRFAQ FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCPKADTLL PVTDTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/g042 95.4% identity in 173 aa overlap

240

	10	20	30	40	50	60
m042-1.pep	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
g042	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m042-1.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
g042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
	130	140	150	160	170	
m042-1.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX					
g042	ATRASLPKIRDVRSICFSPLVRILPLSTVKSVMVVAFFANCSYASAPGPPVMTNCGLWRCR					
	130	140	150	160	170	180
g042	DSQSGSNSVPTVAALSNAAGCKX					
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 173>:

```

a042-1.seq
1   ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTCTCTTTT GATGTATTCG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTGTGC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAACGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA A

```

This corresponds to the amino acid sequence <SEQ ID 174; ORF 042-1.a>:

```

a042-1.pep
1   MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVDSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/a042-1 100.0% identity in 173 aa overlap

	10	20	30	40	50	60
m042-1.pep	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
a042-1	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m042-1.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
a042-1	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
	130	140	150	160	170	
m042-1.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX					
a042-1	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSX					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 175>:

241

```

g043.seq
  1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GTCGGCCCAT CAGCACTTTT
 51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GCCTGAACGC TTCgttgaAC
101  CGTCCCGCGT ggcggtagcc gcAAAAGTGC ATcGCGGCTT GGATGGTGCT
151  GCCCCGATTCG ATGAGGGcga gcGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201  GTCCGGCGAC GGTTCGCGG GTTTGCCTT TGAAATAGCG TTTCAGGTAG
251  CATTCGTTCA GCCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGGCG
301  GGCGAATTCG CTGTTCAAAA TATCGGCGGC TTCGTCTATG CGCCGGCGGC
351  GGTAGCCGTT GTCGTGGCGG CGGAAGGTGA AGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

```

g043.pep
  1  MVVSNQNIYA VGPSALFHIR RQKSVMPER FVEPSRVAVA AKVHRGLDGA
 51  ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQPD AAGDFGDGQRA
101  GEFVQNIIGG FVYAPAAVAV VVAEAGEA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 177>:

```

m043.seq
  1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
 51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101  CGTCCCGCGT GCGGCTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
151  GCCGATTCG ATGAGGGCGA GCGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201  ATCCGGCGAC GGTTCGCGG GTTTGCCTT TGAAATAGCG TTTCAGGTAG
251  CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301  GGCGAATTCG TGTTCAGGA TGTGCGCGC TTCGTCTATG CGCCGACGGC
351  GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

```

m043.pep
  1  MVVSNQNIYA AGPSALLHIR RQKSVMPSE FVEPSRVAVA AKVHGGLDGA
 51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AAGDFGDGQRT
101  GEFVLQDVGG FVYAPTAVTV VVAEAGEAQ*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from *N. gonorrhoeae*:

m043/g043

m043.pep	10	20	30	40	50	60
	MVVSNQNIYAAGPSALLHIRRQKSVMPSE	FVEPSRVAVAAKVHGGLDGAAGFDEGERVF				
g043	MVVSNQNIYAVGPSALFHIRRQKSVMPPE	FVEPSRVAVAAKVHRGLDGAARFDEGERVF				
	10	20	30	40	50	60
m043.pep	70	80	90	100	110	120
	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV					
g043	QPQAAQASGDGFAGLRFEIAFQVAFVQPDAAAGDFGDGQRAGEFAVQNIIGGFVYAPAAVAV					
	70	80	90	100	110	120
m043.pep	130					
	VVAEAGEAQX					
g043	VVAEAGEAXX					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 179>:

```

a043.seq
  1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
 51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101  CGTCCCGCGT GCGGCTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT

```


242

```

151 GCCGGATTTCG ATGAGGGCGA GCGCGTGTTC CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTCGCGG GTTTGCCTT TGAAATAGCG TTTCAGGTAG
251 CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301 GGCGAATTTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```

a043.pep
  1 MVVSNQNIYA AGPSALLHIR RQKSVMPSE FVEPSRVAVA AKVHGGLDGA
 51 AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

m043/a043 100.0% identity in 129 aa overlap

```

              10      20      30      40      50      60
m043.pep    MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              |||||
a043        MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              10      20      30      40      50      60

              70      80      90     100     110     120
m043.pep    QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              |||||
a043        QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              70      80      90     100     110     120

              130
m043.pep    VVAAEGEAQX
              |||||
a043        VVAAEGEAQX
              130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 181>:

```

g044.seq
  1 ATGCTGCCCCG ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTTGA
 51 CCGGCTGTTT GCGGCGGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCGTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTACAGCC TCGATAACG GCGGTCAGCT
201 CCATGCGGTT GTTGGTGGTT TCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 CGGCTGCCGT AGCGCATTA

```

This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>:

```

g044.pep
  1 MLPDQSVEFL PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
 51 GAAAFERFQP FDNGGQLHAV VGGLRFAAEK FFFAAVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 183>:

```

m044.seq
  1 ATGCCGTCCG ACTAGAGCGT CGAGTTCTTT CCAGAAGTCG TCGTTTTTTGA
 51 CGGGCTGTTT GGAGGCGGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCATTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTACAGTC TCGATGACG GCAGTCAGTT
201 CCATGCGGTT GTTGGTGGTT TCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTACCGT AGCGCAyTaa

```

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>:

```

m044.pep
  1 MPSDXSVEFF PEVVVFDGLF GGGFPAVALP TVYPVFHAI F DVLRVGADDD
 51 GAAAFERFQS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 185>:

```
a044.seq
1  GTGCCGTCCG ACCAGCGCGT CGAGTTCTTT CCACAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GGC GGCGGTT TTCCAGCCGT TCGGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTTCAGTCC TTCGATGACG GCGGTCAGTT
201 CCATACGGTT GTTGGTGGTT TCGGCTTCGC CGCCGAAAAG TTCTTTTTTCG
251 TGGCTGCCGT AGCGCATTAA
```

This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

```
a044.pep
1  VPSDQVEFF PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
51  GAAAFERFQS FDDGGQFHTV VGGLRFAAEK FFFVAVAH*
```

m044/a044 91.0% identity over a 89 aa overlap

	10	20	30	40	50	60
m044.pep	MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS					
	: : : : : :					
a044	VPSDQVEFFPQVVVFDGLFGGGFPAVALPTVYPVFHAV FDLRVGADDDGAAAFERFQS					
	10	20	30	40	50	60
	70	80	90			
m044.pep	FDDGSQFHAVVGGLRFAAEKFF FVATVAHX					
	: : :					
a044	FDDGGQFHTVVGGLRFAAEKFF FVAVAHX					
	70	80	90			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from *N. gonorrhoeae*:

m044/g044

	10	20	30	40	50	60
m044.pep	MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS					
	: : : : :					
g044	MLPDQSVEFLPQVVVFDGLFGGGFPAVALPTVYPVFHAV FDLRVGADDDGAAAFERFQP					
	10	20	30	40	50	60
	70	80	90			
m044.pep	FDDGSQFHAVVGGLRFAAEKFF FVATVAHX					
	: : :					
g044	FDNGGQLHAVVGGLRFAAEKFF FAAVAHX					
	70	80	90			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 187>:

```
g046.seq
1  ATGTCGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGcgc gCGCCTGTAT
51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
151 CTGATGGTTT CGGTTATGCC gaATATGGAA AGGCTGCCGt TTTcGTTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TtcgctGGAA CGGACGCGCG
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttC
351 CgcgccggcG Cgttataatg tgAAGGGGGA TGCgcggttg ccgaAAACGG
401 TTTGacatc gaggcggctg CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAAcgg TTACGTCGTT GTTGGTGATG GCGGCAAGGT TTTGCGCGAC
```

244

501 GGTAGAACCT ACCTGCCCCGT TGCCTAAAAT GAGGATTTTC ACGGTATGGG
551 TCGCCGGGTG A

This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:

g046.pep
1 MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLVM AARFCATVEP TCPLPKMRIF TVWVAG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 189>:

m046.seq
1 ATGTCGGCAA TGCTGCGTCC GACAAGCASt CCGC.r.sGC gCGcCTGTAT
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CCGTAACGAG CTGTTCCGGG
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251 CGATCGGGCC GGGGATGTTG AACAGGTCGG CCGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:

m046.pep
1 MSAMLRPTSX PXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 191>:

a046.seq
1 ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CCGTAACGAG CTGTTCCGGG
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251 CGATCGGGCC GGGGATGTTG AACAGGTCGG CCGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:

a046.pep
1 MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*

m046/a046 98.4% identity over a 186 aa overlap

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXXRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					
a046	MSAMLRPTSSPPRRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					

245

	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSLGLRYSRYSLETRAMRPGMLNRSATLQATMLVSSLRESASSKSASSAPA					
a046						
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
a046						
	130	140	150	160	170	180
m046.pep	TVWVAEX					
a046	TVWVAEX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from *N. gonorrhoeae*:

m046/g046

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXXRACMMTIRTRSSAKRKTCNAPGQSIRPASC SVTSCSGLMVSVMPNME					
g046						
	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSLGLRYSRYSLETRAMRPGMLNRSATLQATMLVSSLRESASSKSASSAPA					
g046						
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
g046						
	130	140	150	160	170	180
m046.pep	TVWVAEX					
g046	TVWVAGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 193>:

g047.seq

1	ATGGTCATCA	TACAGGCGcg	gcGCGGCGGG	CTGCTTGTCG	GACGCAGCAT
51	TGCCGACATC	GCCCAAGATT	TGCCCGACGG	GGCCGACTGC	CAAATCTGCG
101	CCGTTTACCG	CAACAACCGC	CTCATCGTCC	CCGCGCCGCA	AACCGTCATC
151	ATCGAAGGCG	ACGAAATCCT	GTTTGCCGCC	GCCGCCGAAA	ACATCGGGGC
201	GGTCATACCC	gaATTGCGCC	CCAAAGAAAC	CAGCACCCGC	CGCATCATGA
251	TTGCCGCGCG	CGGCAACATc	tgctACCGCC	TCGCCAAGCA	GCTCGAACAC

246

```

301 GCATAcaacG TCAAAATCAT CGAATGCCGG CCGCGCcggtg ccgaATGGAT
351 AGCCGAAAAC ctcgAcaaCA CCCTCGTCCT GCAAGGTTCTG Gcaaccgacg
401 aAaccctgct cgAaacgaa tacatcgacg aaatcgaCGT ATTCTGCGCC
451 CTGACCAACG ACGACGAAAG CAACATTAtg tCCGCCCTTT TGGCGAAAAA
501 CCTcggcgCG AAGCgcgtca tcggCATCGT CAACCGCTCA AGCTACGTCG
551 ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601 ACCATCGGCT CGATACTCGC CCACATCCGG CGCGGCGACA TCGTTGCCGT
651 CCACCCCATC CGGCGCGGCA CGGCGGAAGC CATCGAAGTC GTCGCGCACG
701 GCGACAAAAA AACTTCCGCC ATCATCGGCA GGCGCATCAG CGGCATCAAA
751 TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801 AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGTGACCACA
851 TCATCTTTTT CGTCTCGCGC CGGCGCATCC TGAACGAACT GGAGAAACTC
901 ATCCAAGTCA AAATGGGCTT TTTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

g047.pep

```

1 MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDES NIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTS AIIIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 195>:

m047.seq

```

1 ATGGTCATCA TACAGCGcG C..syGCGGA sTGCTTGTCTG GACGCAGCAT
51 TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCC
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCCTCATC
151 ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201 GGTCATACCC GAATTGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GcATCATGAT TkCCGGCGGC GGCAACATCG GCTACCGTCT CGCCAAGCAG
301 CTGGAACACG CATAACAGT yAAAATCATC GAATGCCGCG CGCGCCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCyTG CAAGGTTCTG
401 CAACCGACGA AACCCTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GGCGAaAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCCTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
601 CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
651 CGTTGCCGTC CACCCATCC GGCGCGGCAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCATCAAAT GGCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACCGGCGAA ACCATTATGG GACACCATAC CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTTC GTCTCGCGCC GGCGCATCCT GAACGAAGT
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>:

m047.pep

```

1 MVIIQARXXG XLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETQRN QPXXIMIXGG GNIGYRLAKQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLV LQGSATDETL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGAKRVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAEA IEVVAHGDKK TSAIIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRLNEL
301 EKLIQVKMGF FG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 197>:

a047.seq

```

1 ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCTG GACGCAGCAT

```

```

51  TGCCGACATC  GCCCAAGATT  TGCCCGACGG  GGCCGACTGC  CAAATCTGCG
101 CCGTTTACCG  CAACAACCGC  CTCATCGTCC  CCGCGCCGCA  AACCGTCATC
151 ATCGAAGGCG  ACGAAATCCT  ATTTGCCGCC  GCGCGCGAAA  ACATCGGCGC
201 GGTCATACCC  GAATTGCGCC  CCAAAGAAAC  CAGCACCCGC  CGCATCATGA
251 TTGCCGGCGG  CGGCAACATC  GGCTACCGTC  TCGCCAAGCA  GCTCGAACAC
301 GCATACAACG  TCAAATCAT  CGAATGCCGG  CCGCGCCGTG  CCGAATGGAT
351 AGCCGAAAAC  CTCGACAACA  CCCTCGTCCT  GCAAGGTTTC  GCAACCGACG
401 AAACCTTGCT  CGACAACGAA  TACATCGACG  AAATCGACGT  ATTCTGCGCC
451 CTGACCAACG  ACGACGAAAG  CAACATTATG  TCCGCCCTTT  TGGCGAAAAA
501 CCTCGGCGCG  AAGCGCGTCA  TCGGCATCGT  CAACCGCTCA  AGCTACGTCT
551 ATTTGCTCGA  AGGCAACAAA  ATCGACATCG  TCGTCTCCCC  CCACCTCATC
601 ACCATCGGCT  CGATACTCGC  CCACATCCGG  CGCGGCGACA  TCGTTGCCGT
651 CCACCCCATC  CGGCGCGGCA  CGGCGGAAGC  CATCGAAGTC  GTCGCACACG
701 GCGCAAAAAA  AACTTCCGCC  ATCATCGGCA  GGCGCATCAG  CGGCATCAAA
751 TGGCCCGAAG  GCTGCCACAT  TGCCGCCGTC  GTCCGCGCCG  GAACCGGCGA
801 AACCATTATG  GGACACCATA  CCGAAACCGT  CATCCAAGAC  GCGGACCACA
851 TCATCTTTTT  CGTCTCGCGC  CGGCGCATCC  TGAACGAACT  GGAAAAACTC
901 ATCCAAGTCA  AAATGGGCTT  TTTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

```

a047.pep
1  MVIIQARRGG  LLVGRSIADI  AQDLDPGADC  QICAVYRNNR  LIVPAPQTVI
51  IEGDEILFAA  AAENIGAVIP  ELRPKETSTR  RIMIAGGGNI  GYRLAKQLEH
101 AYNVKIIECR  PRRAEWIAEN  LDNTLVLQGS  ATDETLLDNE  YIDEIDVFCA
151 LTNDDESNI  SALLAKNLGA  KRVIGIVNRS  SYVDLLEGNK  IDIVVSPHLI
201 TIGSILAHIR  RGDIVAVHPI  RRGTAEAIEV  VAHGDKKTS  IIGRRISGIK
251 WPEGCHIAAV  VRAGTGETIM  GHHTETVIQD  GDHIIFFVSR  RRILNELEKL
301 IQVKMGFFG*

```

m047/a047 96.5% identity over a 312 aa overlap

```

              10      20      30      40      50      60
m047.pep      MVIIQARXXGXLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
              | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047           MVIIQARRGGLLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
              10      20      30      40      50      60

              70      80      90      100     110     120
m047.pep      AAENIGAVIPELRPKETQORNQPPXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
              | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047           AAENIGAVIPELRPKETSTR---IMIAGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
              70      80      90      100     110

              130     140     150     160     170     180
m047.pep      AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNI MSALLAKNLGAKRVIGIV
              | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047           AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNI MSALLAKNLGAKRVIGIV
              120     130     140     150     160     170

              190     200     210     220     230     240
m047.pep      NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEIEVVAHGDKK
              | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047           NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEIEVVAHGDKK
              180     190     200     210     220     230

              250     260     270     280     290     300
m047.pep      TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
              | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047           TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
              240     250     260     270     280     290

              310
m047.pep      EKLIQVKMGFFGX
              | | | | | | | | | |

```

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng) from *N. gonorrhoeae*:

m047/g045

m047.pep	MVIIQARXXGXLVGRSIAQDLPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA	60
g047	MVIIQARRGGLLVGRSIAQDLPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA	60
m047.pep	AAENIGAVIPELRPKETQRNQPPXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI	120
g047	AAENIGAVIPELRPKETSTR---IMIAGGGNICYRLAKQLEHAYNVKIIIECRPRRAEWI	117
m047.pep	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	180
g047	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	177
m047.pep	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	240
g047	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	237
m047.pep	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL	300
g047	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL	297
m047.pep	EKLIQVKMGFFGX	313
g047	EKLIQVKMGFFGX	310

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 199>:

g048.seq

1	ATGCTCGACA	AAGGCGAGGA	GTTGCCCGTC	GATTTCACCA	ACCGCTGAT
51	TTACTACGTc	ggcCCcgTCG	ATCCGGTCCG	CGATGAAGTC	GTCCGGTCCCG
101	CAGGTCCGAC	CACAGCCACC	CGCATGGACA	AATTTACCCG	CCAAATGCTC
151	AAACAAACCG	GCCTCTTGGG	CATGATCGGC	AAATCCGagc	gcgGcgcggc
201	cacctGCGAA	GCcatCGCCG	ACAACAAGGC	CGTGTACCTC	ATGGCAGTCG
251	GCGGCGCGGC	ATACCTCGTG	GCAAAAGCCA	TCAAATCTTC	CAAAGTCTTG
301	GCGTTCCTCCG	AATTGGGTAT	GGAAGCCGTT	TACGAATTTG	AAGTCAAAGA
351	TATGCCCGTA	ACCGTCGCCG	TGGACAGCAA	AGGCGAATCC	ATCCACGCCA
401	CCGCCCCGCG	CAAATGGCAG	GCGAAAATCG	GCATCATCCC	CGTCGAGTCT
451	TGA				

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>:

g048 . pep

```

1  MLDKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
51  KQTGLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 AFPELGMEAV YEFEVKMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151 *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 201>:

m048.seq

1	ATGCTCAACA	AAGGCGAAGA	ATTGCCCCGTC	GATTTTACCA	ACCGCCTGAT
51	TTACTACGTC	GGCCCCGTCG	ATCCGGTCCG	CGATGAAGTC	GTCGGTCCGG

```
m048.pep
1  MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
51 EQTDLGLMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151 *
```

```
a048.seq
1 ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCTGAT
51 TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGACGAAATC GTCGGCCCCAG
101 CAGGTCCGAC CACCCGCCACC CGCATGGACA AATTCACCCG CCAAAATGCTC
151 GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGGCGCGGC
201 CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGACCTC ATGGCAGTCG
251 GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301 GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTTG AAGTCAAAGA
351 CATGCCCGTA ACCGTCGCCG TAGACAGCAA AGGCGAATCC ATCCACGCCA
401 CCGCCCCGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
451 TGA
```

```
a048.pep
  1  MLDKGEELPV  DFTNRLIYYV  GPVDPVGVDEI  VGPAGPTTAT  RMDKFTRQML
51  EQTDLLGMIG  KSERGAATCE  AIADNKAVYL  MAVGGAAYLV  AKAIKSSKVL
101 AFPELGMEAI  YEFEVKDMPV  TVAVDSKGES  IHATAPPQWQ  AKIGIIPVKS
151 *
```

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDFTNRLIIYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG					
a048	MLDKGEELPVDFTNRLIIYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m048.pep	KSERGVATCEAIADNKAVYLMVGGAAYLVAKAIAKSSKVLAFPELGMEAIYEFVVKDMPV					
a048	KSERGAATCEAIADNKAVYLMVGGAAYLVAKAIAKSSKVLAFPELGMEAIYEFVVKDMPV					
	70	80	90	100	110	120
	130	140	150			
m048.pep	TVAVDSKGESIHATAPRKWQAKIGIIPVESX					
a048	TVAVDSKGESIHATAPPOWQAKIGIIPVKSX					
	130	140	150			

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF 048.ng) from *N. gonorrhoeae*:

250

m048/g048

	10	20	30	40	50	60
m048 . pep	MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG					
	:					
g048	MLDKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m048 . pep	KSERGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPV					
	:					
g048	KSERGAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAVYEFVVKDMPV					
	70	80	90	100	110	120
	130	140	150			
m048 . pep	TVAVDSKGESIHATAPRKWQAKIGIIPVESX					
g048	TVAVDSKGESIHATAPRKWQAKIGIIPVESX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 205>:

g049 . seq

```

1  ATGCGGGCGC AGGCGTTTGA TCAACCGTTC GGTCAGCTCC TGTTCCGACA
51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
101 TGGACGGGCA TCAACGCTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
151 CCCGTCTGCC GCCGTACCGG ATTCTGCCGC ATCGGCGTTT TCCCCGCCCT
201 CAATCTGTGC GGTTTCAAAT TCGGCACTGT CTTTTTTGGC ATCGAACCGG
251 ATTCTCCGCC GCGATTGAT GTGTTTTTCC GAAAccggca tTTGCAGGGA
301 AGCCTgcgcg TTGAGCCAGT TTTCTGAAG GACGATCATC GGGTCGGTTT
351 CCACTTCCTC GCCGCAATCG GCAACGGCgc tGTTGTGTTC TTCCTGCCAT
401 TTCTTCAGAT ACGCCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 206; ORF 049.ng>:

g049 . pep

```

1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRL FRTAFVFRN
51  PVCRRTGFCR IGVFPALNLC GFKFGTVFFG IEPDSPPRFV VFFRNRHLQG
101 SLRVEPVFLK DDHRVGFDL AAINGAVVF LPPFLQIRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 207>:

m049 . seq (partial)

```

1  ATGCGGGCGC AGGCGTTTGA TCAGCCGTTC GGTCAGCTCC TGTTCCGACA
51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
101 TGGACGGGCA TCAACGTTTC TTCCGCATCG TTTTCCCCGT TTTCCGAAAC
151 CGCCGGCTCA TTCGTGCCGG ATTCTGCCTC GTCGGCGTTT TCCCCGCTTT
201 CAATCTGTCC GGTTTCAAAT TCGACACTGT CTTTTTTGGT ATCAAACCGG
251 ATTCTCCGCC GCGATTGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
301 AGCCTGCGCG TTGAGCCAGT TTTCTGAAG GACGATCATC GGGTCGGTTT
351 CCACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTTC CTCCTGCCAT
401 TTTTTCAGAT ACGCCTT...

```

This corresponds to the amino acid sequence <SEQ ID 208; ORF 049>:

m049 . pep (partial)

```

1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRF FRIVFPVFRN
51  RRLIRAGFCL VGVFPFNLN GFKFDTVFFG IKPDSPPRFV VFFRNRHLQG
101 SLRVEPVFLK DDHRVGFDL AAINGGIVF LLPFFQIRL...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 209>:

a049 . seq

```

1  ATGCGGGGCGC AGGCGTTTGA TCAGCCGTTT GGTGAGCTCC TGTTCCGACA
51  GGCAGAACAC  TTCGCGCCGG TTGACGGCTT TCGGGTTCAG AATATTGATT
101 TGGACGGGCA  TCAACGCTTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
151 CCCGTCTGCC  GCCGTACCCG ATTCTGCCGC ATCGGCGTTT TCCCCGCCTT
201 CAATCTGTCC  GGTTTCAAAT TCGGCACTGT CTTTTTTGGC ATCAAACCGG
251 ATTCTCCGCC  GCGATTTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
301 AGCCTGCGCG  TTGAGCCAGT TTCCTGAAG  GACGATCATC GGGTCGGTTT
351 CGACTTCCTC  GCCGCAATCG GCAACGGCGG CATTGTGTTC CTCCTGCCAT
401 TTTTTCAGAT  ACGCCTT

```

This corresponds to the amino acid sequence <SEQ ID 210; ORF 049.a>:

```

a049.pep
1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLDGHQRF FRTAFAVERN
51  PVCRRTRFCR IGVFPAFNLS GFKEGTVFFG IKPDSPPRFD VFFRNRHLQG
101 SLRVEPVFLK DDHRVGFDL AAIGNGGIVF LLPFFQIRL

```

m049/a049 90.6% identity over a 139 aa overlap

```

              10      20      30      40      50      60
m049.pep    MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRFRRIVFPVFRNRRLIRAGFCL
              |||||:|||||:| |||| | : ||
a049         MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQNIIDLDGHQRFRTAFAVERNPVCRRTRFCR
              10      20      30      40      50      60

              70      80      90      100     110     120
m049.pep    VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFRNRHLQGSRLRVEPVFLKDDHRVGFDL
              :|||||:| |||| | :|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049         IGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFRNRHLQGSRLRVEPVFLKDDHRVGFDL
              70      80      90      100     110     120

              130     139
m049.pep    AAIGNGGIVFLLPFFQIRL
              |||||:| |||| |
a049         AAIGNGGIVFLLPFFQIRL
              130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 049 shows 86.3% identity over a 139 aa overlap with a predicted ORF (ORF 049.ng) from *N. gonorrhoeae*:

```

m049/g049
              10      20      30      40      50      60
m049.pep    MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRFRRIVFPVFRNRRLIRAGFCL
              |||||:|||||:| |||| | : ||
g049         MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRLFRRTAFVFRNPVCRRTGFCR
              10      20      30      40      50      60

              70      80      90      100     110     120
m049.pep    VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFRNRHLQGSRLRVEPVFLKDDHRVGFDL
              :|||||:| |||| | :|||||:|||||:|||||:|||||:|||||:|||||
g049         IGVFPALNLCGFKFGTVFFGIEPDSPPRFDVFFRNRHLQGSRLRVEPVFLKDDHRVGFDL
              70      80      90      100     110     120

              130     139
m049.pep    AAIGNGGIVFLLPFFQIRL
              |||||:| |||| |
g049         AAIGNGAVVFFLPFLQIRLX
              130     140

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 211>:

g050.seq

```

1   atgggCGG GCTGGTGTC TCCCGGCATC TTGGGCATCG GCATCGGCGg
51  cacgcccGAA AAAGccgtgt TGATGGcaaA AGAATCCCTG ATGAGCCACA
101 TCGAcatCca aGaATTGCAG GAAAAAGCCG CGTccggggc ggaattgtcc
151 accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAGGTCA ACGCGCTGGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAT GCCGCCTCCA AACCATTGC CATGATTCCC
301 AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGgACG GCTCAGGtcc
351 TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA

```

This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:

g050.pep

```

1   MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVED*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 213>:

m050.seq

```

1   ATGGGCGCGG GCTGGTGTC TCCCGGCATC TTGGGTATCG GCATCGGCGG
51  C..agCCgAA AAAGCCGTGC TGATGGCAA AGAGTCCCTG ATGAGCCACA
101 TCGACATTCA AGAATTGCAG GAAAAGCCG CGTCCGGCGC GgAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTGGG
201 CATCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA TCCGACCCAC GCCGCCTCCA AACCATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAATC ACGCCGCCG GCGTCGAAGA TGGCCCGATT TGA

```

This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:

m050.pep

```

1   MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDGPI *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 215>:

a050.seq

```

1   ATGGGCGCGG GCTGGTGTC TCCCGGCATC TTGGGCATCG GCATCGGCGG
51  TACGCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
101 TCGACATCCA AGAATTGCAG GAAAAGCCG CGTCCGGCGC GGAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTAGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAC GCCGCCTCCA AACCATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAATC ACGCCGCCG GCGTCGAAGA CTGGCCC

```

This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:

a050.pep

```

1   MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDWP

```

m050/a050 97.7% identity over a 129 aa overlap

```

          10          20          30          40          50          60
m050.pep  MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF
          |||||: |||||: |||||: |||||: |||||: |||||:
a050       MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF

```

253

	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKVNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVEL					
a050	EKVNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
a050	TPPRVEDWP					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from *N. gonorrhoeae*:

m050/g050

	10	20	30	40	50	60
m050.pep	MGAGWCPPGILGIGIGGXAEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF					
g050	MGAGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKVNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVEL					
g050	EKVNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
g050	TPPRVEDX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 217>:

g050-1.seq

```

1  ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGTATCT GCGATGCCTT
51  CCAATTCATC AGCTACTACC ATCCAAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCCGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG GATGTGCAAT
251 GGGATGCGGA CATGAGCGTG GAAAAGATGG TTAACGAAGG CGTACGCCGC
301 GCCTACACTT GGAAGGCAA CACCCTGCGC GCTTCGGTCC TCGCCGATCC
351 GGCCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCACA
401 TGAGCATCGT GCCGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGCTCTG AAAACAAATC CAAACTCGCT ATGCTCAACC CTTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACGATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGCacgc CGAAAAAGCC
601 GTGTTGATGG cgaAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGCGCAAGGC
751 TTGGGCGGTC TGACCACCGT GTTGACGTG AAAATCCTCG ATTACCCGAC
801 CCATGCCGCC TCCAAACCGA TTGCCATGAT TCCCAACTGT GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GTCCTGTCGA ACTCAGCCG
901 CCGCGCGTCG AAGACTGACC CGATCTGACT TACAGCCCCG ACAACGGCAA
951 ACGCGTCGAT GTCGATAAGC TGACCAAAGA AGAAGTGGA AGCTGGAATA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCGATGCC
1051 GCGCAACAA GCCTCGTCAA TATGCTCGAC AAAGGCGAGG AGTTGCCCGT
1101 CGATTTTACC AACCGCCTGA TTTACTACGT CGGCCCGTGC GATCCGGTCG
1151 GCGATGAAGT CGTCGGTCCC GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTTATCC GCCAAATGCT CAAACAAACC GGCCTCTTGG GCATGATCGG

```

```

1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAGG
1301 CCGTGTACCT CATGGCAGTC GGCGGCGCGG CATACCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGTA TGGAGCCGT
1401 TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGTCGCC GTGGACAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGC GCAAATGGCA GCGGAAATC
1501 GGCATCATCC CCGTCGAGTC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:

g050-1.pap

```

1  MTVIKQEDFI  QSICDAFOFI  SYYPKDYID  ALYKAWQKEE  NPAAKDAMTQ
51  ILVNSRMCAE  NNRPICQDTG  IATVFLKVG  DVQWDADMSV  EKMVNEGVRR
101  AYTWEQNTLR  ASVLADPAGK  RQNTKDNT  VIHMSIVPGG  KVEVTCAAKG
151  GGSENKSKLA  MLNPSDNIV  WVLKTIPT  AGWCPPGILG  IGIGGTPEKA
201  VLMAKESLMS  HIDIQELQEK  AASGAELST  EALRLELF  VNALGIGAQG
251  LGGLTTVLDV  KILDYPTHA  SKPIAMIP  AATRHVEFEL  DSGGPVELTP
301  PRVED*PDLT  YSPDNKRVD  VDKLTKEE  SWKTGDVLL  NGKILTGRDA
351  AHKRLVNMLD  KGEELPVDF  NRIIYVGP  DPVGDEVVGP  AGPTTATRMD
401  KFTROMLKQT  GLLGMIGK  RGAATCEA  DNKAVYLM  GGAAYLVAKA
451  IKSSKVLAFP  ELGMEAVY  EVKDMPTV  VDSKGESI  TAPRKWQAKI
501  GIIPVES*

```

g050-1/p14407

```

sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
>gi|280063|pir|B44511 fumarate hydratase (EC 4.2.1.2) fumB, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
Score = 172 bits (432), Expect = 4e-42
Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)

```

```

Query: 11  QSICDAFOFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENNRPICQDTG 70
Q+ DA + H K L+ E + K Q L NS + A+ P CQDTG
Sbjct: 53  QAFHDASFMLRPAHQKQVAAAILHDPEASEND---KYVALQFLRNSEIAAKGVLPCTCQDTG 109

Query: 71  IATVFLKVGMDVQWDADMSVEKMVNEGVRRAYTWEQNTLRASVLADPAGKRQNTKDNTPA 130
A + K G V W E+ +++GV Y E N + A K NT N PA
Sbjct: 110  TAIIVGKKGQRV-WTGGGD-EETLSKGVNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166

Query: 131  VIHMSIVPGGKVEVTCAAKGGGSENKSKL-----AMNPSDNIVDWVLKTIPTMGAGWCP 185
I + V G + + C AKGGG NK+ L A+L P + +++++ + T+G CP
Sbjct: 167  QIDLAVDGDGEYKFLCVAKGGGSANKTYLYQETKALLTPG-KLKNFLVEKMRTLGTAAACP 225

Query: 186  PXXXXXXXXXTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFKVNXXX 245
P T + L + +H EL + + L EL E+
Sbjct: 226  PYHIAFVIGGTS AETNLKTVKLASAHY-YDELPTGNEHGQAFRDVQLEQELLEEAQKLG 284

Query: 246  XXXXXXXXXXXTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSG----PVELTPP 301
D++++ P H AS P+ M +C+A R+++ +++ G +E P
Sbjct: 285  LGAQFGGKYFAH-DIRVIRLPRHGASCPVGMGVSCSADRNIKAKINREGIWEKLEHNPG 343

Query: 302  RVEDXPDLTYSPDNKRVDVDKLTKE--EVASWKTGDVLLNGKILTGRDAAHKRLVNM 358
+ +VD+++ KE +++ + L L G I+ GRD AH +L +
Sbjct: 344  QYIPQELRQAGEGEAVKVDLNRPMKEILQLSQYPVSTRLSLTGTIIVGRDIAHAKLKL 403

Query: 359  LDKGEELPVDFNRLIYXXXXXXXXXXXXXXXXTTATRMDKFTROMLKQTGLLGMIGK 418
+D G+ELP + IYY TTA RMD + + G + M+ K
Sbjct: 404  IDAGKELPQYIKDHPIYYAGPAKTAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463

Query: 419  SERGAATCEAIADNKAVYLMVAVGG-AAYLVAKAIKSSKVLAFPELGMEAVYEFVVKDMPV 477
R +A + YL ++GG AA L ++IK + +A+PELGMEA+++ EV+D P
Sbjct: 464  GNRSQQVTDACHKHGGFYLSIGGPAAVLAQQSIKHLECVAYPELGMEAIWKIEVEDFPA 523

Query: 478  TVAVDSKG 485
+ VD KG
Sbjct: 524  FILVDDKG 531

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 219>:

m050-1.seq

```

1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
51  CCAATTTCATC AGCTACTATC ATCCCAAAGA CTACATCGAC GCGCTTTATA
101  AGGCGTGGCA GAAGGAAGAA AATCCTGCCG CCAAAGACGC GATGACGCAG
151  ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACACCGGCC CCATCTGCCA
201  AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG AACGTCCAAT
251  GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC

```

```
301 GCCTACACTT GGGGAAGGCAA TACGCTGCGC GCTTCCGTCC TCGCCGATCC
351 GGCCCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCATA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGCTCTG AAAACAAATC CAAACTCGCC ATGCTCAATC CTTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGT ATCGGCATCG GCGGCACGCC CGAAAAAGCC
601 GTGCTGATGG CAAAAGAGTC CCTGATGAGC CACATCGACA TTCAAGAATT
651 GCAGGAAAAA GCCCGCTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGCACAAGGC
751 TTGGGCGGAC TGACCACCGT GTTGACGTG AAAATCCTCG ATTATCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAACTGC GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCAGCCG
901 CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCGC ACAACGGCAA
951 ACGCGTCGAT GTCGACAAGC TGACCAAGA AGAAGTGGCA AGCTGGAAAA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCAAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTACC AACCGCCTGA TTACTACGT CGGCCCGTTC GATCCGGTCG
1151 GCGATGAAGT CGTCGGTCCG GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAACC GACCTCTTGG GCATGATCGG
1251 CAAATCCGAG CGCGGCGTGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301 CCGTGTACCT CATGGCAGTC GCGGCGCGCG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGTTTCCCC GAATTGGGCA TGGAAAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGATAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGCG GCAAATGGCA GGCGAAAATC
1501 GGCATCATCC CCGTCGAATC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>:

m050-1.pep

```
1 MTVIKQEDFI QSIDAFQFI SYHPKDYID ALYKAWQKEE NPAAKDAMTQ
51 ILVNSRMCAE NNRPICQDTG IATVFLKVG M NVQWDADMSV EEMVNEGVRR
101 AYTWEQNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151 GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELFEK VNALGIGAQG
251 LGGLTTVLVD KILDYPTHA SKPIAMIPNC AATRHVEFEL DSGSPVELTP
301 PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351 AHKRLVDM LN KGEELPVDFE NRIYYVGPV DPVGDEVVGP AGPTTATRMD
401 KFTROMLEQT DLLGMIGKSE RGVATCEAIA DNKAVYLM AV GGAAYLVAKA
451 IKSSKVLAF P ELGMEAIYEF EVKDMPVTVA VDSKGESIHA TAPRKWQAKI
501 GIIPVES*
```

m050-1/g050-1 98.2% identity in 507 aa overlap

```
10 20 30 40 50 60
m050-1.pep MTVIKQEDFIQSIDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
|||||
g050-1 MTVIKQEDFIQSIDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
10 20 30 40 50 60

70 80 90 100 110 120
m050-1.pep NNRPICQDTGIATVFLKVG MNVQWDADMSVEEMVNEGVRRAYTWEQNTLRASVLADPAGK
|||||
g050-1 NNRPICQDTGIATVFLKVGMDVQWDADMSVEKMNVEGVRRAYTWEQNTLRASVLADPAGK
70 80 90 100 110 120

130 140 150 160 170 180
m050-1.pep RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
|||||
g050-1 RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
130 140 150 160 170 180

190 200 210 220 230 240
m050-1.pep AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK
|||||
g050-1 AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK
190 200 210 220 230 240

250 260 270 280 290 300
m050-1.pep VNALGIGAQGLGGLTTVLVDKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP
|||||
g050-1 VNALGIGAQGLGGLTTVLVDKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP
250 260 270 280 290 300
```

	310	320	330	340	350	360
m050-1.pep	PRVEDWPDLTYS PDNGKRVDVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLN					
g050-1	PRVEDXPDLTYS PDNGKRVDVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVNMLD					
	310	320	330	340	350	360
	370	380	390	400	410	420
m050-1.pep	KGEELPVDFTNRLIYYVGPVDPVGVDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE					
g050-1	KGEELPVDFTNRLIYYVGPVDPVGVDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIGKSE					
	370	380	390	400	410	420
	430	440	450	460	470	480
m050-1.pep	RGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPVTVA					
g050-1	RGAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAVYEFVVKDMPVTVA					
	430	440	450	460	470	480
	490	500				
m050-1.pep	VDSKGESIHATAPRKWQAKIGIIPVESX					
g050-1	VDSKGESIHATAPRKWQAKIGIIPVESX					
	490	500				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 221>:

```
a050-1.seq
1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
51  CCAATTCATC AGCTACTACC ATCCCAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AACCCGCGCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGCAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
201 AGATACCGGT ATCGCGACCG TGTTTTGGAA AGTCGGTATG GATGTGCAAT
251 GGGATGCAGA CATGAGCGTC GAAGAGATGG TTAACGAAGG CGTGCGCCGC
301 GCCTACACTT GGAAGGCAA TACGCTGCGC GCTTCCGTTT TCGCCGACCC
351 CGCCGGCAAA CGCCAAAATA CCAAAGACAA CACGCCCGCC GTCATCCATA
401 TGAGCATCGT GCCGGCGGAC AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GCGGTTCTTG AAAACAAATC CAACTCGCC ATGCTCAACC CTTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATTCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGTACGCC CGAAAAGCC
601 GTGTTGATGG CGAAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCACGCGC TAGGCATCGG CGCGCAAGGC
751 TTGGGCGGTC TGACCACCGT GTTGACGTG AAAATCCTCG ATTACCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAAGTGC GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCACGCCG
901 CCGCGCGTCG AAGACTGGCC CGATTGACT TACAGCCCGG ACAACGGCAA
951 ACGGCTCGAT GTCGACAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCAG CCGCGATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCGAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCGGTC GATCCGGTCG
1151 GCGACGAAAT CGTCGGCCCA GCAGGTCCGA CCACCGCCAC CCGCATGGAC
1201 AAATTCAACC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCTCGCC GACAACAAAG
1301 CCGTGTACCT CATGGCAGTC GGCGGCGCGG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGCA TGGAAAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGACAGCA
1451 AAGGCGAATC CATCCAGGCC ACCGCCCGG CCCAATGGCA GGCAGAAATC
1501 GGCATCATCC CCGTCAAATC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>:

```
a050-1.pep
1  MTVIKQEDFI QSIDAFQFI SYYHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM DVQWDADMSV EEMVNEGVR
101 AYTWEENTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGD KVEVTCAAKG
151 GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGAQG
251 LGGLTTVLDV KILDYPTHA SKPIAMIPNC AATRHVEFEL DSGSPVELTP
301 PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351 AHKRLVDMLD KGEELPVDFT NRIYYVGPV DPGDEIVGP AGPTTATRMD
401 KFTROMLEQT DLLGMIGKSE RGAATCEAIA DNKAVYLMV GGAAYLVAKA
451 IKSSKVLAF ELGMEAIYEF EVKDMPVTVA VDSKGESIHA TAPPQWQAKI
501 GIIPVKS*
```

a050-1/m050-1 98.4% identity in 507 aa overlap

```

a050-1.pep      10      20      30      40      50      60
MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
m050-1          10      20      30      40      50      60
MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE

a050-1.pep      70      80      90      100     110     120
NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTRLASVLADPAGK
m050-1          70      80      90      100     110     120
NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTRLASVLADPAGK

a050-1.pep     130     140     150     160     170     180
RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
m050-1         130     140     150     160     170     180
RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG

a050-1.pep     190     200     210     220     230     240
AGWCPPGILGIGGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
m050-1         190     200     210     220     230     240
AGWCPPGILGIGGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK

a050-1.pep     250     260     270     280     290     300
VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP
m050-1         250     260     270     280     290     300
VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP

a050-1.pep     310     320     330     340     350     360
PRVEDWPDLTYSFDNGKRVVDVKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLD
m050-1         310     320     330     340     350     360
PRVEDWPDLTYSFDNGKRVVDVKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLN

a050-1.pep     370     380     390     400     410     420
KGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
m050-1         370     380     390     400     410     420
KGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE

a050-1.pep     430     440     450     460     470     480
RGAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPVTVA
m050-1         430     440     450     460     470     480
RGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPVTVA

a050-1.pep     490     500
VDSKGESIHATAPPQWQAKIGIIPVKSX
m050-1         490     500
VDSKGESIHATAPRKWQAKIGIIPVESX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 223>:

```

g052.seq
1   ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GGCGGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CCACTCCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351 CGCCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 224; ORF 052.ng>:

```
g052.pep
```



```

      1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
     51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
    101 RLRLETTWSP ACRKVKNAA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 225>:

```

m052.seq
      1 ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTCAAAGG
     51 CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
    101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
    151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
    201 GGC GGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACTCCGA
    251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
    301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
    351 CGCCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 226; ORF 052>:

```

m052.pep
      1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
     51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
    101 RLRLETTWSP ACRKVKNAA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 227>:

```

a052.seq
      1 ATGGCTTTGG TCGCGGAGGA AACGGAAATA TCCGCGCCGT GTTCAAAGG
     51 CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
    101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCTCCC
    151 AAGGGATTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
    201 GGC GGCTTTC CATTCTGTTA TATCAGTCGG CGACACGTGA CTCACTTCGA
    251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
    301 AGGCTGCGGC TGGAAATCAC ATGGTCGCCC GCCTGCAAAA AGGTGAAAAA
    351 CGCCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 228; ORF 052.a>:

```

a052.pep
      1 MALVAEETEI SAPCFKG*EP TGDSRLLSTT KSAPMPCANS AKASKSATSP
     51 KGLDGVSKNS SLVLALTAAF HSFISVGD* LTSMPNLVTM LLIKPTVVPN
    101 RLRLEITWSP ACKKVKNAA*

```

m052/a052 95.8% identity over a 119 aa overlap

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
a052	MALVAEETEI SAPCFKGX EPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m052.pep	SLVLALTA AFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLETTWSPACRKVKNAAX					
a052	SLVLALTA AFHSFISVGD TXLTSMPNLVTMLLIKPTVVPNRLRLEITWSPACKVKNAAX					
	70	80	90	100	110	120

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 052 shows 95.8% identity over a 119 aa overlap with a predicted ORF (ORF 052.ng) from *N. gonorrhoeae*:

m052/g052

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					

```

g052      MALVAETEISAPCFKGCEPTGDSRLLSTTKSAPMPCANSKASKSATSPKGLDGVSKNS
           10          20          30          40          50          60
           70          80          90          100         110         120
m052.pep  SLVLALTAAFHSFISVGDITWLTSMPLNLTMLLIKPTVVPNRLRLLEITWSPACKKVKNAX
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g052      SLVLALTAAFHSFISVGDITRLTPMPLNLTMLLIKPTVVPNRLRLLETTWSPACRKVKKNAX
           70          80          90          100         110         120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 229>:

```

g073.seq
1  ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
51  TCCGCCGATG CCGTCTGAAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCAGCCGA CAGGCGCACC
151 AGTCCGGGGC GGatacCGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAAAAGAGT TCCACGACTT TCCACGCGGC TGCTTGGTCTG
301 GCGACTTCAA AACCAGATGAC GATGCCGCCG CCGTTTGTCT GTTTGCGGAT
351 AAGCTCCGCC TGCGGATGGT CGGGCAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>:

```

g073.pep
1  MCMPYAIRVS DGICRIFPPM PSETRNQRAS ACFKSSIKSP TYSKPTDRRT
51  SPGRIPAASF SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAWS
101 ATSKPMTMPP PFCCLRISSA CGWSGNPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 231>:

```

m073.seq
1  ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
51  GCCGTCTGAA ACACGCAATC AGCGTGCGAG TGCCTGTTTC AAATCGTCAA
101 TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
151 CGGATGTTGG CGGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGGT
201 TGTCACGGG TGGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
251 GGGAAAAGAG TTCCACGCCG TCCACAACCT TCCACGCCGC TTCTTGATCG
301 GCAACTTCAA AGCCGATGAC GATGCCGCCG CCGTTTGTCT GTTTGCGGAT
351 AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

```

m073.pep
1  MCMPYKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTYS KPTDRRTNPG
51  RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSSTP STTFHAASXS
101 ATSKPMTMPP PFCCLRISAA XGWSNPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 233>:

```

a073.seq
1  ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
51  TCCGCCGATG CCGTCTGAA. CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCAGCCGA CAGGCGCACC
151 AATCCGGGGC GGATGTTGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT
301 GCTTGGTCTG CCACTTCAAA GCCGATGACG ATGCCGCCGC CGTTTGTCTG
351 TTTGCGGATA AGCTCCGCCT GAGGATGGTC GGGTAATCCG GTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>:

```

a073.pep
1  TCMSYKIRVS DGICGVFPPM PSEXRNQRAS ACFKSSIKSP TYSKPTDRRT
51  NPGRMILAASF SSGCILPCVV VHGWVVERT SPRLAVREKS STPSTTFHAA
101 AWSATSKPMT MPPPFCLRI SSA*GWSGNP V*

```

m073/a073 92.3% identity over a 130 aa overlap

260

	10	20	30	40	50
m073.pep	MCMPIKIRVSDGICCC--PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF				
a073	TCMSYKIRVSDGICGVFPMPSEXRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF				
	10	20	30	40	50
	60	70	80	90	100
m073.pep	SSGCILPCVVHGWVVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPFFCCLRI				
a073	SSGCILPCVVHGWVVERTSPRLAVREKSSTPSTTFHAAAWSATSKPMTMPPFFCCLRI				
	70	80	90	100	110
	120	129			
m073.pep	SAAXGWSNDPVX				
a073	SSAXGWSGNPVX				
		130			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from *N. gonorrhoeae*:

m073/g073

	10	20	30	40	50
m073.pep	MCMPIKIRVSDGICCC--PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF				
g073	MCMPIAIRVSDGICRIFPPMPSETRNQRASACFKSSIKSPTYSKPTDRRTSPGRIPAASF				
	10	20	30	40	50
	60	70	80	90	100
m073.pep	SSGCILPCVVHGWVVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPFFCCLRI				
g073	SSGCILPCVVHGLVMVERTSPRLAVREKSST--TFHAAAWSATSKPMTMPPFFCCLRI				
	70	80	90	100	110
	120	129			
m073.pep	SAAXGWSNDPVX				
g073	SSACGWSGNPVX				
		120			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 235>:

g075.seq

```

1  ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAAAATA CAAAAAGCGC
51  GCGGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 GCGCTTCCAA AGCGTTTTTTT GCCGTTTCGG GCAACGCTGC GTTTCCTGT
151 GCGGCCAAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTT ACGAAATTTT
251 TAAAAAATG TGTTTGCGGG CTTTGTGAAG GTTTTAGAGA CCGCTGCCG
301 GGCCTCTTAA ACTTAATCTT CTTTTTCGTA GAATCCGAAA ATTACAAATT
351 CCGCGCCTAT CTCTTCCAAT GCCGAGCTAA AAGCGTCTTC ATAGCTGTCA
401 TATTTACCGG CTGA

```

This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>:

g075.pep

```

1  MPPYFITLLT MENTKSAKT PTTIQPASIP SAFAASKAFF AVSGNAAFAC
51  AAKASGAATV TASFAPYLRQ VLINFMIFSF TKFLKKCVCG LCEGFRDRLP
101 GLLNLIFFFV ESENYKFPAY LFQCRKSVF IAVIFTG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 237>:

```
m075.seq
  1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAATA CAAAAAGCGC
 51  GCGGAAAATG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101  CGGCTTCCAA AGCGTTTTTT GCGGTATCGG GCAACGTTGC ATTTGCATGT
151  GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201  TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251  TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301  TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351  CTTCCAAACC TCGGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401  TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

```
m075.pep
  1  MPSYFITLLT MENTKSAAKM PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51  AAKARGAAVT TASFAPYLRL VLNFMIFS FKKCLAVMDGA FFRPPNIRK
101  SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from *N. gonorrhoeae*:

```
m075/g075

      10      20      30      40      50      60
m075.pep  MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
          ||||||||||||||||| ||||||||||||||||| |||||||:||||| |||||
g075      MPPYFITLLTMENTKSAAKPTTIQPASIPSAFAASKAFFAVSGNAAFACAAKASGAAVT
          10      20      30      40      50      60

      70      80      90      100     110
m075.pep  TASFAPYLRLQVLINFMIFS-----KKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVAD
          ||||||||||||||||| |||: : :| | |: : :| | : : :| | : : :|
g075      TASFAPYLRLQVLINFMIFSFTKFLKKVCGLCEGFRDRLPGLLNLIFFFESENYKFPAY
          70      80      90      100     110     120

      120     130
m075.pep  FFQTCVNRFFEVEEIIIGIGDX
          :|| : :| :| :|
g075      LFQCRAKSVFIIVFTGX
          130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 239>:

```
a075.seq
  1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAAGA CAAAAAGCGC
 51  GCGGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101  CGGCTTCCAA AGCGTTTTTT GCTGTATCGG GCAACGTTGC ATTTGCATGT
151  GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201  TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251  TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301  TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351  CTTCCAAACC TCGGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401  TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>:

```
a075.pep
  1  MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51  AAKARGAAVT TASFAPYLRL VLNFMIFS FKKCLAVMDGA FFRPPNIRK
101  SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

m075/a075 98.5% identity over a 136 aa overlap

```

      10      20      30      40      50      60
m075.pep MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      MPSYFITLLTMEKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          10      20      30      40      50      60

      70      80      90     100     110     120
m075.pep TASFAPYLQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      TASFAPYLQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          70      80      90     100     110     120

      130
m075.pep CVNRFFEVEIIGIGDX
          |||||
a075      CVNRFFEVEIIGIGDX
          130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 241>:

```

g080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTGCG TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTTGCA CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCGTTTTG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACgaGA TGAAACGCCT CCgGCTTTTT
601 ACCgAAGCGT GGCAGCATCT gttgcGTAAG AATAAAATC GGTTATCCTA
651 TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
701 GTTTACCCGA AAAAGAAATcc gAAGAAAtatt gggaacaggt ttgggacata
751 ttacggcctg gcgtcggaaa cggttcgacg caaatttcaa tcagttatAA
801 GGCAGAcga acaatggaac AGcagtaa

```

This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:

```

g080.pep
1  MWDNAEAMER LTRWLLVMMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDIVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RGPMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
251 LRPGVNGGST QISISYKGRR TMEQQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 243>:

```

m080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTGCG TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTTCGCG CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT

```

601 ACCGAAGCGT GGCAGCATCT GTTGCCTAAA AATAAAAATC GGTATCCTA
 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG
 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>:

m080.pep
 1 MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLVPK QVSLKGNLVY
 51 SDKKTLGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDVE
 101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
 151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng) from *N. gonorrhoeae*:

m080/g080

	10	20	30	40	50	60
m080.pep	MWDNAEAMERLTRWLLVMMA MLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKTLGSLA					
080	MWDNAEAMERLTRWLLVMMA MLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKALGSLA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m080.pep	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDVEVVLTERKPVARWGDHALVDG					
080	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDVEVVLTERKPVARWGDHALVDG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m080.pep	EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
080	EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWNVVL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m080.pep	DNGITVRLGRENEMKRLRLFTEAWQHLLRK NKNRLSYVDMRYKDGFSVRYASDGLPEKES					
080	DNGITVRLGRENEMKRLRLFTEAWQHLLRK NKNRLSYVDMRYKDGFSVPHAPDGLPEKES					
	190	200	210	220	230	240
m080.pep	EEX					
080	EEYWEQVWDILRPGVGNGSTQISISYKGRRTMEQQX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 245>:

a080.seq
 1 ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
 51 CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
 101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTAGTTTAT
 151 TCCGATAAGA AAGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
 201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
 251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
 301 GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
 351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGTTTGGAC AGACCCGGAA
 401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
 451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
 501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA

```

551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
601 ACCGAAGCGT GGCAACATCT GTTGCCTAAA AATAAAAATC GGTATCCTA
651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTCCCGACG
701 GTTTACCCGA AAAAGAATCC GAAGAATAG

```

This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:

a080.pep

```

1 MWDNAEAMER LTRWLLVMM MLLAASGLVW FYNSNHLVPK QVSLKGNLVY
51 SDKKALGSLA KEYIHGNI LR TDINGAQEAY RRYPWIASVM VRRRFPDVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RGPMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK KNRLSYVDM RYKDGFSVRY APDGLPEKES EE*

```

m080/a080 99.2% identity over a 242 aa overlap

	10	20	30	40	50	60
m080.pep	MWDNAEAMERL LTRWLLVMMAMLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKTGSLA					
a080	MWDNAEAMERL LTRWLLVMMAMLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKTGSLA					
	10	20	30	40	50	60
m080.pep	KEYIHGNI LR TDINGAQEAY RRYPWIASVM VRRRFPDVE VVLTERKPVARWGDHALVDG					
a080	KEYIHGNI LR TDINGAQEAY RRYPWIASVM VRRRFPDVE VVLTERKPVARWGDHALVDG					
	70	80	90	100	110	120
m080.pep	KEYIHGNI LR TDINGAQEAY RRYPWIASVM VRRRFPDVE VVLTERKPVARWGDHALVDG					
a080	KEYIHGNI LR TDINGAQEAY RRYPWIASVM VRRRFPDVE VVLTERKPVARWGDHALVDG					
	70	80	90	100	110	120
m080.pep	EGNVFEARLD RGPMPVFRGA EGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
a080	EGNVFEARLD RGPMPVFRGA EGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
	130	140	150	160	170	180
m080.pep	EGNVFEARLD RGPMPVFRGA EGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
a080	EGNVFEARLD RGPMPVFRGA EGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
	130	140	150	160	170	180
m080.pep	DNGITVRLGR ENEMKRLRLF TEAWQHLLRKKNRLSYVDMRYKDGFSVRYASDGLPEKES					
a080	DNGITVRLGR ENEMKRLRLF TEAWQHLLRKKNRLSYVDMRYKDGFSVRYAPDGLPEKES					
	190	200	210	220	230	240
m080.pep	DNGITVRLGR ENEMKRLRLF TEAWQHLLRKKNRLSYVDMRYKDGFSVRYASDGLPEKES					
a080	DNGITVRLGR ENEMKRLRLF TEAWQHLLRKKNRLSYVDMRYKDGFSVRYAPDGLPEKES					
	190	200	210	220	230	240
m080.pep	EEX					
a080	EEX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 247>:

g081.seq

```

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AACAAACCCG TGTGCGCAT CGTAACCGAC AGCCGCGATA
101 TTCGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACCGG
151 CATGATTTTG TTGGAGGCGT ATTGTCTGCG GCGCGGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGCGCGCA TAATGTGAAC
301 CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CCGTGAAGGA
351 GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG
401 CGACGGCAGG CAACTTCAAC AACCACAtcg gaTTGCCGCT GACTTTATTG
451 AAATtaaaAcg aAAAAACCCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTTGGcgaa ctggcgggtTt taacgcaaaT CGCCAAACCC GATGCCGCTT
551 TGGTcaACAA CGCCCTGCGC GCCCATGTCTG GATGCGGTTt cgacggagt9
601 GCGGATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTCAGA
651 CCGCATGGCA CTGATTCTCT AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CCGCAACGTT TAATTTGAAT ACGTGCACTT TCGGCGTCGA TAGCGGCGAT
751 GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTTGATTT

```

```

801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
951 AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG CATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAGGAc gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGcggaAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGAGCC ACGATTTGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT GCGCTTTATG CAGAtggaAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTga

```

This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:

g081.pep

```

1 MKPLDLNFIC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
51 HDFVGGVLSA GAAAVVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNALR AHVGCDFDGV
201 GDIKAKKSEI YAGLCSDGMA LIPQEDANMA VFKTATFNLN TCTFGVDSGD
251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNA AAAAAALAAG
301 LSLNDVAEGL QGFSNIKGR LNVKAGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGD MGELGED EAAAMHAEVG AYARDQIEA AYFVGDN SVE
401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 249>:

m081.seq

```

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGAGCG GTTTGACGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC ATTGCAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
301 CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CCGTGAAGGA
351 AATGCTGGCT GCGGTATTGC GCCgCCGTTT CGGCGATGAT GCCGTGTTGG
401 CGACGGCAGG CAACTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
451 AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTCCGCGAA CTGGCGGTTT TAACGCAMAT CGCCAAACCA AATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTG GCTGCGGTTT CGACGGAGTG
601 GGCGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTCAGA
651 CGGCATTGCA CTGATTCCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTTACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATT
801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCGGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCCT GACAGCATGA AAGCTGCGAT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAACTG GCGGAGGACG AAGCCGCGCG TATGCACGCC GAAGTCGGCG
1151 CGTATGCCCG CGACCAAGGC ATCGAAGCGG CTTATTTTGT CGGCGACAAC
1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCCGCCG
1251 CAAAGACCCG TTGATTCAAG TGTTGCGCCA CGATTTGCCC GAACGCGCCA
1301 CCGTGTGGT GAAAGGTTG CGCTTTATGC AGATGGAAGA AGTGGTCGAG
1351 GCATTGGAGG ATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:

m081.pep

```

1 MKPLDLNFIC QALKLPMPSE SKPVSRIVTD SRDIRAGDVF FALAGERFDA
51 HDFVEDVLAA GAAAVVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTXIAKP NAALVNNAMR AHVGCDFDGV

```


Homology with a predicted ORF from *N. gonorrhoeae*

m081/g081

	10	20	30	40	50	60
m081.pep	MKPLDLNFICQALKLPMPSESKPVSRIVTDSRDIRAGDVFFALAGERFDAHDFVEDVLAA					
g081	MKPLDLNFICQALKLPMPSENKPVSRIVTDSRDIRREGDVFFALAGGRFDAHDFVGGVLSA					
	10	20	30	40	50	60
m081.pep	70	80	90	100	110	120
	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITSGSGKTTVKEMLA					
g081	GAAAVVVSREDCAALGGALKVDDTLAALQTLAKAWRDNVNPFVFGITSGSGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	130	140	150	160	170	180
	AVLRRRFGDDAVLATAGNFNNHIGLPLTLKLNEKHRYAVIEMGMNHFGELAVLTXIAKP					
g081	AVLRRRFGDDAVSATAGNFNNHIGLPLTLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP					
	130	140	150	160	170	180
m081.pep	190	200	210	220	230	240
	NAALVNNAMRAHVGC GFDGVGDI AKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN					
g081	DAALVNNALRAHVGC GFDGVGDI AKAKSEIYAGLCSDGMALIPQEDANMAVFKTATFNLN					
	190	200	210	220	230	240
m081.pep	250	260	270	280	290	300
	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVVLVPVGRHNVHNAAAAALALAAG					
g081	TCTFGVDSGDVRAENIVLKPLSCEFDLVCGDERTAVVLVPVGRHNVHNAAAAALALAAG					
	250	260	270	280	290	300
m081.pep	310	320	330	340	350	360
	LSLNDVAEGLKGFSNIKGRNLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV					
g081	LSLNDVAEGLQGFSNIKGRNLNVKAGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV					
	310	320	330	340	350	360
m081.pep	370	380	390	400	410	420
	MGDMGELGELGEDEAAAMHAEVGAYARDQGIEAAYFVGDNVSVEAAEKFGADGLWFAAKDP					
g081	MGDMGELGE--DEAAAMHAEVGAYARDQGIEAAYFVGDNVSVEAAEKFGADGLWFAAKDP					
	370	380	390	400	410	
m081.pep	430	440	450			
	LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX					
g081	LIQVLSHDLPERATVLVKGSRFMQMEEVVEALEDKX					
	420	430	440	450		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 251>:

a081.seq

```

1  ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51  GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGATGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCG GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGCGTTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTACCGC GTTGCAAATG TTGGCGAAGG CGTGGCGCGA GAATGTGAAC
301 CCGTTTGTGT TCGGTATTAC CGGCTCGGGC GGCAAGACGA CGGTGAAGGA
351 AATGTTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATAAT GCCGTTTGG
401 CGACGGCAGG CAACTTCAAC AACCACATCG GATTGCCGTT GACTTTGTTG
451 AAATTAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GTATGAACCA
501 TTTTGGCGAA CTGGCGGTTT TGACACAAAT CGCCAAACCC GATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCT GCTGCCGTTT CGACGGAGTG
601 GGCATATTG CCAAAGCGAA AAGCGAGATT TATCAAGGCT TATGTTTCTG
651 CGGCATGGCG CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTCCACGCGG AAAATATCGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
801 GGTGTGCGGC AACGAGTGCG CAGCCGTGGT TCTGCCCGTT CCCGCCGCC
851 ACAATGTCCA CAACGCCGCC GCCGCCGCCG CGCTGTCTTT GGCTGCAGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCCT GACAGCATGA AAGCTGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCCGCCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGTGAGGAC GAAGCCGCCG CCATGCACGC CGAAGTCGGC GCGTACGCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGCGGAAA AATTTGGCGC GGACGGTTTG TGTTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGCGCC ACGATTTGCC CGAACGCCCC ACCGTGTTGG
1301 TGAAGGTTT CCGCTTTATG CAGATGGAAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>:

a081.pep

```

1  MKPLDLNFIC QALKLMPSE SKPVSRIVTD SRDIRAGDVF FALAGGRFDA
51  HDFVEDVLAA GAAVVVSRE DCVAMDGALK VDDTLTALQM LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDN AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNAMR AHVGCDFDGV
201 GDIKAKSEI YQGLCSDGMA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251 VHAENIVLKP LSCEFDLVCG NECAAVVLPV PGRHNVHNAA AAAALS LAAG
301 LSLNDVAEGL KGFSNIKGR LNVKSGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQIEA AYFVGDNSVE
401 AAEKFGADGL WFAAKDPLIQ VLRHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

m081/a081 96.7% identity over a 455 aa overlap

```

              10      20      30      40      50      60
m081.pep      MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGERFDAHDFVEDVLAA
              |||
a081           MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGGRFDAHDFVEDVLAA
              10      20      30      40      50      60

              70      80      90      100     110     120
m081.pep      GAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA
              |||
a081           GAAVVVSREDCVAMDGALKVDDTLTALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA
              70      80      90      100     110     120

              130     140     150     160     170     180
m081.pep      AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGE LAVLTQIAKP
              |||
a081           AVLRRRFGDNAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGE LAVLTQIAKP

```

268

	130	140	150	160	170	180
	190	200	210	220	230	240
m081.pep	NAALVNNAMRAHVGC GFDGVGDI AKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN					
a081	DAALVNNAMRAHVGC GFDGVGDI AKAKSEIYQGLCSDGMALIPQEDANMAVFKTATLNLN					
	190	200	210	220	230	240
	250	260	270	280	290	300
m081.pep	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVLPVPGRHNHVNAAAAAALALAAG					
a081	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGNECAAVLPVPGRHNHVNAAAAAALSLAAG					
	250	260	270	280	290	300
	310	320	330	340	350	360
m081.pep	LSLNDVAEGLKGF SNIKGR LNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV					
a081	LSLNDVAEGLKGF SNIKGR LNVKSGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV					
	310	320	330	340	350	360
	370	380	390	400	410	420
m081.pep	MGDMGELGELGEDEAAAMHAEV GAYARDQGIEAAYFVGDN SVEAAEKFGADGLWFAAKDP					
a081	MGDMGELGE---DEAAAMHAEV GAYARDQGIEAAYFVGDN SVEAAEKFGADGLWFAAKDP					
	370	380	390	400	410	
	430	440	450			
m081.pep	LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX					
a081	LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX					
	420	430	440	450		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 253>:

```

g082.seq
1  aTGTGGTTGT TGAAGTTGCC TGCCGTCGCC GAAACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATCTC CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAACGGGT TCACATTATC GCGCCACGCC
151 TTCGCCAACG TTTGCAACGC GGCAAGCGTG TCATCGACTT TCAACGCGCC
201 GCCCAAAGCC GCGCAATCTT CGCGCGAAAC CACAACGCC GCCGCGCCG
251 CAGACAATAC GCCTCCAACA AAATCATGCG CGTCAAACCG CCCGCCGCC
301 AATGCGAAAA ACACATCGCC TTCCCGAATA TCGCGGCTGT CGGTTACGAT
351 GCGCGACACG GGTGTGTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401 AGATGAAATT TAGGTCCAGT GGTTTCATAT TTGCTTTCGT TAATATTCGG
451 GCGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGCATT TTTCTGTACG
551 TATCATTTTT TAGACGTATT TTTAGCCGAT TTGCCTTTT CCGCATACCA
601 CGGCGCGGGG TCGTCGGACT GTCTGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCCGGCAC ATCGGGGACA TTCCCCAAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTCGATACC CGTCCAATG CCGAATCCGC GTAA

```

This corresponds to the amino acid sequence <SEQ ID 254; ORF 082.ng>:

```

g082.pep
1  MWLLKLPAVA ETASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTLSRHA
51  FANVCNAASV SSTFNAPPKA AQSSRETTA AAPADNTPPT KSCASNRPPA
101 NAKNTSPSRI SRLSVTMRDT GLFSDGIGSL RAWQMKFRSS GFIFAFVNIR
151 AADTSVAADF FIACFAVVKH RLFSHSHSAF FLYVSFFRRI FSRFAFSRIP
201 RRGVGLSVD KGVIAFARH IGDIPKIIA VIGQLVGFD T RPTAESA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 255>:

```

m082.seq
1  ATGnnGTTGT TGAAGTTGCC TGCCGTCGCC AACACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATTTT CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAACGGAT TCACATTTT ACGCCACGCC
151 TTTGCCAGCG TTTGCAATGC GGCAAGCGTG TCATCGACTT TCAACGCGCC

```

```

201 ATCCATTGCA GCACAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCAG
251 CAGCCAATAC GTCTTCAACA AAATCATGCG CGTCAAACCG CTCGCCCCGCC
301 AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTACGAT
351 GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401 AGATGAAATT TAGGTCCAGT GGTTTCATAT TTAATTTTCGT TAATATTCGG
451 GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTA AACAC AGATTATTTT CCCATTCTCA TTCGGsATTT TTTCTGTACG
551 TATCATTTTT TAGACGTATT TTAGTCGAT TTGCCTTTTC CCGCATACCA
601 CGGCGCGGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCCTGCAC ATCGGGAACA TTCCCCAAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTTGATACC CGTCCAACG CCGAATCCGC GTAA

```

This corresponds to the amino acid sequence <SEQ ID 256; ORF 082>:

```

m082.pep
  1 MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEPVM PN TNGFTFSRHA
 51 FASVCNAASV SSTFNAPSIA AQSSRETTTA AAPAANTSST KSCASNRSPA
101 NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVVKH RLFSHSHSXF FLYVSFFRRI FSRFAFSRIP
201 RRGVVGQSVD KGKVIAFALH IGNIPPKIIA VIGQLVGFD T RPTAESAX

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from *N. gonorrhoeae*:

```

m082/g082

      10      20      30      40      50      60
m082.pep MXLLKLPAVANTASSPKRRRNTAASISFTVVLPPEPVMNTNGFTFSRHAFASVCNAASV
          | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g082      MWLLKLPAVAETASSPKRRRNTAASISFTVVLPPEPVMNTNGFTLSRHAFANVCNAASV
          | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      10      20      30      40      50      60

      70      80      90     100     110     120
m082.pep SSTFNAPSIAAQSSRETTTAAAPANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g082      SSTFNAPPKAAQSSRETTTAAAPADNTPPTKSCASNRPPANAKNTSPSRISRLSVTMRDT
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      70      80      90     100     110     120

      130     140     150     160     170     180
m082.pep GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF
          | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g082      GLFSDGIGSLRAWQMKFRSSGFIFAFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      130     140     150     160     170     180

      190     200     210     220     230     240
m082.pep FLYVSFFRRIFSRFAFSRIPRRGVVQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFD T
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g082      FLYVSFFRRIFSRFAFSRIPRRGVVGLSVDKGKVIAFARHIGDIPPKIIAVIGQLVGFD T
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      190     200     210     220     230     240

m082.pep RPTAESAX
          | | | | | |
g082      RPTAESAX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 257>:

```

a082.seq
  1 ATGTGGTTGT TGAAGTTGCC TGCCGTCGCC AAAACGGCAT TATCGCCGAA
 51 ACGGCGGCGC AATACCGCAG CCAACATTC CTTACCGTC GTCTTGCCGC

```

a082.pwp

m082/a082 95.5% identity over a 247 aa overlap

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 259>:

```
g084.seq
1  ATGAAacaAT  CCGcccgaat  aAAAAATATG  GATCAGACAT  TAAAAAATAc
51  attgggcatt  tGCGCGcttt  tagcctTTTG  TTTTggcgcg  gccATCGCAT
101 CAGGTTATCA  CTGGAATAT  GAATACGGCT  ACCGTATTTC  TGCGTGGGCG
151 GCTTTGGCTT  CGGTTGTATT  TTTATTATTA  TTGGCACGCG  GCTTCCCGCG
201 CGTTTCTTCA  GTTGTTTTAC  TGATTTACGT  CGGCACAACC  GCCCTATATT
251 TGCCGGTCCG  CTGGCTGTAT  GGTGCGCCTT  CTTATCAGAT  AGTCGGTTTC
301 ATATTGAAA  CCAATCTGAC  CGAGGCGCGT  GAATTGTTCG  GCAATCTTCG
351 CGGGTCGCTT  TATTTTGTGC  AGGCATTATT  TTTCATTTTT  GGCTTGACAG
```

g084 . pep

m084.seq

m084 . pep

m084/q084

	10	20	30	40	50	
m084 . pep	MKQSARIKXMNQTLTYLGICALLTF-----YHPEYEYGYRYSAVGALASVVFLLL					
	: :					
g084	MKQSARIKNMDQTLKNTLIGICALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL					
	10	20	30	40	50	60
	60	70	80	90	100	110
m084 . pep	LARGFPRVSSVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL					
g084	LARGFPRVSSVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL					
	70	80	90	100	110	120
	120	130	140	150	160	170
m084 . pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTIILLTILSCAVMDKIASDKDL					
g084	YFVQALFFIFGLTVWKYCVSVGVFADVKNYKRRSKIWLTIILLTILSCAVMEKIAGDKDW					
	130	140	150	160	170	180

	180	190	200	210	220
m084 . pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX				
g084	REPDAGLLLNIFDLYYDLAFRAGTICRQARPHFGSSKKSVMAYPPTCAQVX				
	190	200	210	220	230

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 263>:

a084 . seq

```

1  ATGAAACAAT CCGCCCGAAT AAAAAATATG GATCAGACAT TAAAAAATAC
51  ATTGGGCATT TGC GCGCTTT TAGCCTTTTG TTTTGGCGCG GCCATCGCAT
101 CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGT
151 GCTTTGGCTT CCGTTGTATT TTTATTATTA TTGGCACGCG GTTTCGCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251 TGCCGTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301 ATATTGGAAT GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351 CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCAATTTT GGCTTGACAG
401 TTTGGAGATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTG
501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTTC CGAGAACCTG
551 ATGCCGGCCT GTTGTGTAAT ATTTTCGACC TGTATTACGA TTTGGCTTCC
601 .GCGCCGGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA

```

This corresponds to the amino acid sequence <SEQ ID 264; ORF 084.a>:

a084 . pep

```

1  MKQSARIKNM DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VLLIYVGT ALYLPVGWLY GAPSYQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWRYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLN IFDLYYDLAS
201 XAGTICRQAR PHFGSSKKSVMAYPSCCAQ V*

```

m084/a084 92.2% identity over a 231 aa overlap

	10	20	30	40	50	60
m084 . pep	MKQSARIKXMNQTLTYTLGICALLTFXXXXXXHYHPEYGYRYSAVGALASVVFLLL					
	: : :					
a084	MKQSARIKNMDQTLKNTLGIALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL					
	10	20	30	40	50	60
m084 . pep	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL					
a084	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL					
	70	80	90	100	110	120
m084 . pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
	: : :					
a084	YFVQALFFIFGLTVWRYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
	130	140	150	160	170	180
m084 . pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
a084	REPDAGLLLNIFDLYYDLASXAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
	190	200	210	220	230	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 265>:

g085 . seq

```

1  ATGGGCAAAG GGCAGGACTT CACGCCCTG CGCGACGCGT TGAAAGATAA

```

```

51  GGCAAAAGGC GTGTTCTGA TCGGCGTCGA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTTGGAAGAG
151 GCGGTTTCAGA CGGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACC GTT
251 CGGAAGTGTT tatCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:

g085.pep

```

1  MGKGQDFTPL RDALKDKAKG VFLIGVDAPO IRRDLGCGCL NLTDCVTLEE
51 AVQTAYAQA E AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 267>:

m085.seq

```

1  ATGGGTAAAG GGCAGGACTT CACGCCCTG CGCGATGCAC TGGTAGGCAA
51 GGCAAAAGGC GTGTTCTTGA TTGGTGTGCA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTTGGGAGAA
151 GCGGTTTCAGA CGGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTGCT
201 CAGCCCCGCC TCGCGAGCT TTGATATGTT CAAAGGCTAC GCGCACC GTT
251 CGGAAGTGTT TATCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:

m085.pep

```

1  MGKGQDFTPL RDALVGKAKG VFLIGVDAPO IRRDLGCGCL NMTDCATLGE
51 AVQTAYAQA E AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng) from *N. gonorrhoeae*:

m085/g085

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGCLNMTDCATLGEAVQTAYAQA E					
g085	MGKGQDFTPLRDALKDKAKGVFLIGVDAPOIRRDLGCGCLNLTDCVTLEEAVQTAYAQA E					
	10	20	30	40	50	60
	70	80	90			
m085.pep	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
g085	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
	70	80	90			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 269>:

a085.seq

```

1  ATGGGCAAAG GGCAGGACTT CACGCCCTG CGCGACGCGC TGCCGGCAA
51 GGCAAAAGGC GTGTTCTGA TCGGTGTCGA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTTGGAAGAA
151 GCGGTTTCAGA AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACC GTT
251 CGGAAGTGTT TATCGGGGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:

a085.pep

```

1  MGKGQDFTPL RDALAGKAKG VFLIGVDAPO IRRDLGCDL NMTDCATLEE
51 AVQKAYAQA E AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

m085/a085 94.7% identity over a 94 aa overlap

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGCLNMTDCATLGEAVQTAYAQA E					


```

|||||
a085  MGKGQDFTPLRDALAGKAKGVFLIGVDAFQIRRDLDGCDLNMTDCATLEEAVQKAYAQAEE
      10      20      30      40      50      60

      70      80      90
m085.pep AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
|||||
a085  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
      70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 271>:

```

g086.seq
1  ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTG GCATCGAAGG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GGCAGGCGGG GTTCGTCGTT GCCGGCCTTA TAGCGAGCGG TTTTATATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTCG
201 CTTATCCGGC CTGTTGCTGG TAGCCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTACACGC
351 CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
401 GCGGGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAAACAT TAGAAATGTA CGGcCGTTTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTTC
551 GTTCGTTTGT CGTCATTACC GTCATTACCG TTGGAATGCT GTTCTGCGCA
601 GGATTGCCGT GGAAATATTT TTTTGTCTCG GTAGGCAGCG TCTTGGGTGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTGGGA CCCGTGGAAA GACCCGACAG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGCAGATTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
901 ATATTCGTGT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGtttgA CTTTCAACGC CTATATCGCT TCGGTATCG
1001 GCATTTGGAT CCGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 CTTTGGCGA CCAAAGTCT GACGctgCcg tTGATGTCCT ATGGcggTTC
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATCGATT
1151 ATGAAAACCG CCAGAAAATG CGCGGTTACC GGGTGGAGTA AA

```

This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:

```

g086.pep
1  MVVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGFLW
51  FLCMRTRWR LVPWIFALSG LLLVAVLIAG REINGATRWI PLGPLNFQPT
101 ELFKLAVILY LASLFTRREE VLRSMSLWQ QSIWRGTANL IMSATNPQAR
151 RETLEMYGRF RAILLPIMLV AFGLVLIMVQ PDFGSFVVIT VITVGMFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFDPWK DPQGAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301 IFCYGWLVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SFFNIGVNIG
351 ALPTKGLTLP LMSYGGSSVF FMLISMLLL RIDYENRQKM RGYRVE*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 273>:

```

m086.seq
1  ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTG GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTCG
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTCA AGCtGGCGGT CATCCTTTAT TTGGCAAGCC TGTTACACGC
351 CCGTGAAGAA GTGTTGcGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
401 GCGGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGrCACGT
451 CGTGAAACAT TAGAAATGTA CGGCCGTwTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTTC

```

```

551 GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTGGCA
601 GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTGGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGCAGATTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGAT CGGkTCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA mCAAAGyCT GACGCGGCCG Tg .AtGTCCw ATGGCGGTTT
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTkG CGTATAGATT
1151 ATGAAAACCG CCGGAAAATG CGCGGTTATC GGGTGGAGTA A

```

This corresponds to the amino acid sequence <SEQ ID 274; ORF 086>:

```

m086 . pep
  1  MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
 51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
101  ELFKLAVILY LASLFTRREE VLRSMESLWQ QSIWRGTANL IMSATNPQXR
151  RETLEMYGRX RAILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
201  GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
251  HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301  IFCYGLVVR AFSIGQSRD LGLTFNAYIA SGIGIWIGXQ SFFNIGVNIG
351  ALPXKGLTXP XMSXGGSSVF FMLISMMLLX RIDYENRRKM RGYRVE*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 086 shows 96.7% identity over a 396 aa overlap with a predicted ORF (ORF 086.ng) from *N. gonorrhoeae*:

m086/g086

	10	20	30	40	50	60
m086 . pep	MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
g086	MVVLMTAFGLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
	10	20	30	40	50	60
m086 . pep	LVPWIFALSGLLLVVLIAGREINGATRWIPLGPLNFQPTLFLKAVILYLASLFTRREE					
g086	LVPWIFALSGLLLVAVLIAGREINGATRWIPLGPLNFQPTLFLKAVILYLASLFTRREE					
	70	80	90	100	110	120
m086 . pep	VLRSMESLWQSIWRGTANLIMSATNPQXRRETLEMYGRXRRAILPIMLVAFGLVLIMVQ					
g086	VLRSMESLWQSIWRGTANLIMSATNPQARRETLEMYGRFRRAILPIMLVAFGLVLIMVQ					
	130	140	150	160	170	180
m086 . pep	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFDPWK					
g086	PDFGSFVVITVITVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFDPWK					
	190	200	210	220	230	240
m086 . pep	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
g086	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
	250	260	270	280	290	300
m086 . pep	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
g086	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
	310	320	330	340	350	360

a086.seq

1	ATGGTGGTGC	TGATGACGGC	GTTCAGCCTG	CTGATGATTT	ATTCGGCTTC
51	TGTGTATTTG	GCATCAAAAG	AAGGCGGCGA	TCAGTTTTTC	TATTTGACCA
101	GACAGCGCGG	GTTTCGTCTG	GCCGGCTTGA	TAGCGAGCCG	TTTGTATGCG
151	TTTCTTTGCA	GGATGAGGAC	ATGGCGGCGG	CTTGTGCCGT	GGATTTTTGC
201	CCTATCCGGC	CTGTTGCTGG	TAGTCGTATT	GATTGCCGGG	CGCGAAATCA
251	ATGGCGCGAC	CCGTTGGATA	CCTTTGGGTC	CGTTGAATTT	CCAGCCGACC
301	GAGCTGTTCA	AGCTGGCGGT	CATCCTTTAT	TTGGCAAGCC	TGTTCAACGG
351	CCGTGAAGAA	GTGTTGCGCA	GCATGGAAG	TTTGGGTTGG	CAGTCGATTT
401	GGCGGGGGAC	GGCCAATCTG	ATCATCTCC	CCAGCAATCG	GCAGGCAGCT
451	CGTGAACAT	TAGAAATGTA	CGGCCGTTTC	CGGGCGATCA	TCCTGCCGAT
501	TATGCTGGTG	GCGTTCGGTT	TGGTGCTGAT	AATGGTACAG	CCGGATTTCT
551	GTTTCGTTGT	CGTCATTACC	GTCATTGCCG	TTGGAATGCT	GTTTTTGGCA
601	GGATTGCCGT	GGAAATATTT	TTTCGTCCTG	GTAGGCAGCG	TCTTGGGCGG
651	GATGGTGCTG	ATGATTACCG	CCGTCCTCTA	CCGTGTGCAG	CGGGTAGTGG
701	CATTTTTTGA	CCCGTGGAAA	GACCCGCAGG	GTGCCGGCTA	CCAGCTTACC
751	CACTCTCTGA	TGGCAATCGG	GCGCGGAGAG	TGGTTCGGTA	TGGGTTTGGG
801	TGCGAGTTTG	AGCAAACGCG	GCTTTCTGCC	GGAAGCGCAT	ACCGATTTTA
851	TTTTTGCCAT	CATCGCCGAA	GAATTCGGTT	TCTTCGGTAT	GTGCTGTCTG
901	ATATPTCTGT	ACCGCTGGCT	GGTGGTGCGG	GCGTTTTCCA	TCGGCAAGCA
951	GTCGCGCGAT	TTGGGTTTGA	CTTTCAACGC	CTATATCGCT	TCGGGTATCG
1001	GCATTTGGAT	CGGTATCCAA	AGTTTCTTCA	ATATCGGTGT	GAACATCGGT
1051	GCTTTGCCGA	CCAAAGGCTC	GACGCTGCCG	TTGATGTCTC	ATGGCGGTTT
1101	GTCAGTCTTT	TTCATGCTGA	TACAGATGAT	GCTGCTGTTG	CGTATAGATT
1151	ATGAAAACCG	CCGGAATATG	CGCGGTTACC	GGGTGGAGTA	A

a086.pcp

1	MVVLMTAFSL	LMIYSASVYL	ASKEGGDQFF	YLTRQAGFVV	AGLIASGLLW
51	FLCRMRTWRR	LVPWIFALSG	LLLVLVVLIAG	REINGATRWI	PLGPLNFQPF
101	ELFKLAVIYR	LASLFTREE	VLRSMESLGW	QSIWRGTANL	IMSATNPQAR
151	RETLEVMYGRF	RAIILPIMLV	AFGLVLIMVQ	PDFGSFVVIT	VIAVGMLFLA
201	GLPWKYFFVL	VGSVLGGMVL	MITAAPYRVQ	RVVAFLDPWK	DPQGAGYQLT
251	HSLMAIGRGE	WFGMGLGASL	SKRGFLPEAH	TDFIFAIIE	EFEGFEGMVL
301	IFCYGILVVR	AFSICQSRD	LGLTFNAYIA	SGIGIWIQI	SFFNIGCVNI
351	ALPTKGLTLF	LMSYGGSSVF	FMLISMLLL	RIDYENRRKM	RGYRVE*

	10	20	30	40	50	60
m086.pep	MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
a086	MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m086.pep	LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLFLKLAIVILYASLFTREE					
a086	LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLFLKLAIVILYASLFTREE					

	70	80	90	100	110	120
	130	140	150	160	170	180
m086.pep	VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLMVQ					
a086	VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLMVQ					
	190	200	210	220	230	240
m086.pep	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
a086	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	250	260	270	280	290	300
m086.pep	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAI AEEFGFFGMCVL					
a086	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAI AEEFGFFGMCVL					
	310	320	330	340	350	360
m086.pep	IFCYGWLVRVAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXKGLTXP					
a086	IFCYGWLVRVAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP					
	370	380	390			
m086.pep	XMSXGGSSVFFMLISMMLLXRIDYENRRKMRGYRVEX					
a086	LMSYGGSSVFFMLISMMLLLRIDYENRRKMRGYRVEX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 277>:

g087.seq

```

1  ATGGGCGGTA AAACCTTTAT GCTGATGGCG GCGGGAACGG GCGGACACAT
51  TTTCCAGCT CTGGCTGTGG CGGATTCATT GCGCGTGCGC GGTCATCATG
101 TAATTTGGCT GGGCAGCAAG GATTTCGATG AAGAGCGCAT CGTGCCGCAA
151 TACGGCATA GCTTGGAAAC GCTGGCGATT AAAGGAATAC GCGGCAACGG
201 CATCAAACGC AAGCTGATGC TTCCGTTTAC TCTGTACAAA ACCGTCCGCG
251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
301 GCGGTTTTTG TTACCTTTCC CGGCGGTCTG GCGGCGAAAC TCTTGGGCGT
351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGCTTG TCCAACCGCC
401 AccTGTCGCG ctGGGCGAAA CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
451 AGCCACGAAG GCGGTTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGCGAAGGC CGTCTGAAAA
551 TTTTGGTGGT CGGCGGCAGT TTGGGTGCGG ACGTTTGTAA CAAAACCGTA
601 CCGCAGGCGT TGGCACTGCT GCCTGAAGAG GTGCGCCCGC AGATGTACCA
651 CCAGTCGGGG CGTAACAAGC TGGGCAATCT TCAGGCGGAT TATGACGCGT
701 TGGGCGTGAA AGCGGAATGC GTGGAATTTA TTACCGACAT GGTGTCCGCC
751 TACCGTGATG CCGATTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC
801 CGAGTTGACG GCGGCGGGG TGGGCGCGTT GTTAGTGCCG TATCCTCACG
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTCATGGT GCAGGCAGAA
901 GCGGGGCTGC TGTGCGGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
951 AATCCTCGGC AGCCTCAACC GCGAAAAATG CCTCAAATGG GCGGAAAACG
1001 CCCGTACGTT GGCATTGCCG CACAGCGCGG ATGACGTTGC CGAAGCCGCG
1051 ATTGCGTGTG CGGCGTAAA

```

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>:

g087.pep

```

1  MGGKTFMLMA GGTGGHIFPA LAVADSLRVR GHHVIWLGSK DSMEERIVPQ
51  YGIRLETIAI KGIRNGIKR KLMLPFTLYK TVREAQRIIR KHRVECVIGF
101 GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVLNKTV

```

201 PQALALLPEE VRPQMYHQSG RNKLGNLQAD YDALGVKAEC VEFITDMVSA
 251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
 301 AGLLLPQTQL TAEKLAELILG SLNREKCLKW AENARTLALP HSADDVAEEA
 351 IACAA*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 279>:

m087.seq
 1 ATGGGCGGTA AAACCTTTAT GCTGawkkCG GCGGGAACGG GCGGACATAT
 51 TTTCCCCGCG CTGGCGGTGG CCGATTCAAT GCGCGCGCGC GGCCATCATG
 101 TGATTTGGCT GGGCAGCAAG GATTGATGG AAGAGCGTAT CGTGCCGCAA
 151 TACGGCATAAC GCTTGGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
 201 CATCAAACGC AAACGTATGC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG
 251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCCT CATCGGCTTC
 301 GCGGCTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTArGCGT
 351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
 401 ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
 551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
 601 CCGCATGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC ATATGTACCA
 651 CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGnnn nnnnnnnnnn
 701 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 751 nnnGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG CGGTTGACGA
 801 TCACCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG GCGGGATTGC
 851 TGTGTCGCA AACCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC
 901 GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT
 951 GGCAGTCCG CACAGTGCGG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG
 1001 CGGCGTAA

This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:

m087.pep
 1 MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHVHVIWLGSK DSMEERIVPQ
 51 YGIRLETLAI KGVRGNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF
 101 GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLSRWAK RVLVAFPKAF
 151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVLNKT
 201 PHALALLPDN ARPHMYHQSG RGKLGILQAX XXXXXXXXXXXX XXXXXXXXXXXX
 251 XAGLGALLVP YPHAVDDHQT ANARFMVQAE AGLLLPQTQL TAEKLAELILG
 301 GLNREKCLKW AENARTLALP HSADDVAEEA IACAA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng) from *N. gonorrhoeae*:

m087/g087

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVHVIWLGSKDSMEERIVPQYGIRLETLAI					
	: : : : :					
g087	MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHHVHVIWLGSKDSMEERIVPQYGIRLETLAI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m087.pep	KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGFVTFPGGLAAKLLXVPIV					
	: : : : :					
g087	KGIRGNGIKRKLMLPFTLYKTVREAQRIIRKHRVECVIGFGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
	: : : : :					
g087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
	130	140	150	160	170	180

279

	190	200	210	220	229
m087.pep	RLKILVVGSLGADVLNKTVP	PHALALLPDNARPHMYHQS	GRGKLGI	LQA-----	
g087	RLKILVVGSLGADVLNKTVP	QALALLPEEVRPQMYHQS	GRNKLGNLQADYDALGV	KAEC	
	190	200	210	220	230 240
			230	240	250
m087.pep	-----		AGLGALLVPYPH	AVDDHQTANARFMVQAE	
g087	VEFITDMVSAYRDADLVIC	RAGALTIAELTAAGL	GALLVPYPH	AVDDHQTANARFMVQAE	
	250	260	270	280	290 300
	260	270	280	290	300 310
m087.pep	AGLLLPQTQLTAEKLA	EILGGLNREKCLKWA	ENARTLALPHS	ADDVAEAAIACAAX	
g087	AGLLLPQTQLTAEKLA	EILGSLNREKCLKWA	ENARTLALPHS	ADDVAEAAIACAAX	
	310	320	330	340	350

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 281>:

a087.seq

```

1  ATGGGCGGTA AAACCTTTAT GCTGATGGCG GCGGGAACGG GCGGACATAT
51  TTTCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
101 TAATTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA
151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGCGTGC GCGGCAACGG
201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG
251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
301 GCGGCTTCG TTACCTTTCC CGGCGGTTG GCGGCGAAGT TATTAGGCGT
351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTT TCCAACCGCC
401 ACCTGTGCGC CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCGC ATATTAGCAA
501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTTGA CAAAACCGTA
601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA
651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC
701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC
751 TACCGCGATG CCGATTGTTG GATTTGCCGT GCCGGCGCGC TGACGATTGC
801 CGAGTTGACG GCGGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG
901 GCGGGATTGC TGTTGCCGCA AACCAGTTG ACGGCGGAAA AACTCGCGGA
951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG
1001 CCCGTACGTT GGCAGTCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG
1051 ATTGCGTGTG CGGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 282; ORF 087.a>:

a087.pep

```

1  MGGKTFMLMA GGTGGHIFPA LAVADSLRAR GHVHVLGSK DSMEERIVPQ
51  YDILLETLAI KGVNRNGIKR KLMLPFTLYQ TVREAQIIIR KHRVECVIGF
101 GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLVAFPKAF
151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVLNKTVP
201 PQALALLPDN ARPQMYHQS RGKLGSLQAD YDALGVQAE VEFITDMVSA
251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
301 AGLLLPQTQL TAEKLAELG GLNREKCLKW AENARTLALP HSADDVAEAA
351 IACAA*

```

m087/a087 85.4% identity over a 355 aa overlap

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGH	HHVHVLGSKDSMEERIVPQY	GIRLET	LAI		
a087	MGGKTFMLMAGGTGGHIFPALAVADSLRARGH	HHVHVLGSKDSMEERIVPQY	DILLET	LAI		
	10	20	30	40	50	60

280

	70	80	90	100	110	120
m087.pep	KGVNRNGIKRKLMPLVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPFIV					
a087	KGVNRNGIKRKLMPLFTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
a087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m087.pep	RLKILVVGGSGLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQAXXXXXXXXXXX					
a087	RLKILVVGGSGLGADVLNKTVPQALALLPDNARPMYHQSGRGKLSLQADYDALGVQAE					
	190	200	210	220	230	240
	250	260	270	280	290	300
m087.pep	XX-----XXXXXXXXXAGLGALLVPYPHAVDDHQTANARFMVQAE					
a087	VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE					
	250	260	270	280	290	300
	290	300	310	320	330	
m087.pep	AGLLLPQQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
a087	AGLLLPQQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 283>:

g088.seq

1	ATGTTT	TAT	GGCTCG	CACA	TTTCAG	CAAC	TGTTAA	ACCG	GTCTGA	AATAT	
51	TTTTC	AATAC	ACCAC	ATTCC	GCGCCG	TAT	GGCGCG	GTTG	ACCGC	CTTGG	
101	CGTTT	TCCCT	GATGT	TCGGC	CCGTGG	ACGA	TACGCA	GGCT	GACCG	CGCTC	
151	AAATG	CGGGC	AGGCAG	TGCG	TACCGA	CGGC	CCGCAA	ACCC	ACCTCG	TCAA	
201	AAACG	GCACG	CCGACG	ATGG	GCGGTT	CGCT	GATTCT	GACC	GCCATT	ACCG	
251	TGTCC	ACCCT	GTTGT	GGGGC	AACTGG	GCGA	ACCGT	TATAT	CTGGAT	TCTC	
301	TTGGG	CGTAC	TGCTTG	CCAC	CGGTGC	GCTC	GGTTTT	TACG	ACGACT	TGCG	
351	CAAA	AGTC	GTT	TATA	AAGACC	CCAACG	GCGT	GTCCGC	CAAA	TTCAAA	TGG
401	TGTGG	CAGTC	AAGCG	TGCC	GTTatcgc	CG	GTtggca	TT	GTTTT	TACctt	
451	gCcg	CAATT	CCGCCA	ACAA	TATTTT	GATT	GTCCCG	tTtT	TCAA	ACAAAT	
501	CGCC	TGCCG	CTGGG	CGTGG	TGGCT	Tttt	gGtg	ttgTCT	TACCT	GACCA	
551	TCGT	CGGCAC	ATCCA	ACGCC	GTCA	ACCTCA	Cga	ACGGCTT	GGAC	GCCTT	
601	GCCG	Ccttcc	cgttc	gtcct	cgttg	ccgcC	GGGCT	CGCCA	ttttc	gcctA	
651	CGTC	AGCGGA	CACTA	CCAAT	TTTCCA	ATA	CCTC	CAGCTT	CCCT	ATGTCG	
701	CCGG	CGCGAA	CGAAG	TCGCT	ATATT	CTGCA	CCGC	CATGTG	CGGC	CGTGC	
751	CTCG	GATTTT	TGTGG	TCAA	CGCCT	TATCC	GCGCA	AGTCT	TTAT	TGGCGA	
801	TGTC	GGCGCG	CTGGC	ATTGG	GTGCC	CGCGCT	CGGT	accGtt	gCCG	TcaTcg	
851	tCCG	CCAAGA	ATTTG	Tcctc	gtcatta	TGG	GCGGT	CTGTT	cgtc	gtagaa	
901	gccgtg	TCCG	TTATG	CTCa	tgtcgg	CTGG	TACA	AAGAAA	Ccaaaaa	acg	
951	CATCT	TcCTg	acgG	caccga	ttcatc	acca	ttaC	Caactt	cgatg	CTGGa	
1001	aagaa	acgca	agtcg	tcgtc	CGTTt	CTGGA	TTAt	Taccat	cgtcg	tggtt	
1051	tTgat	aggtt	tGagt	accct	caAA	attcgc	ggaa	actatg	ccgT	CCGAAC	
1101	ACCTT	TCAGA	CGGC	ATTGA	ACGCG	CAATA	A				

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>:

g088.pep

1	MFLW	LAHFSN	WLTGL	NIFQY	TTFRA	VMAAL	TALAF	SLMFG	PWTIR	RRLTAL
51	KCGQ	AVRTDG	PQTHL	VKNGT	PTMGG	SLILT	AITVS	TLLWG	NWAN	PYIWIL
101	LGVL	LATGAL	GFYDD	WRKV	YKDPN	GVSAK	FKMVW	QSSVA	VIAG	LALFYL
151	AANS	ANNILI	VPFFK	QIALP	LGVVG	FLVLS	YLTIV	GTSNA	VNLT	DGLDGL

m088.seq

1	ATGTTTAT	GGCTCGACA	TTTCAGCA	TGTTAACCG	GTCTGAAT
51	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
101	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
151	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
201	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
251	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
301	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
351	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
401	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
451	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
501	nnnnnnnnnn	nnnnGGCGTGG	TCGGCTTTT	GGTGTGTCT	TACCTGACCA
551	TCGTCGGCAC	ATCCAATGCC	GTCAACCTCA	CCGACGGCTT	GGACGGCCTT
601	GCGACCTTCC	CCGTCGTCCT	CGTTGCCGCC	GGCCTCGCCA	TCCTCGCCTA
651	TGCCAGCGGC	CACTCACAAT	TTGCCCAATA	CCTGCAATTA	CCTTACGCTG
701	CCGGCGCAAA	CGAAGTGGTG	ATTTTCTGTA	CCGCCATGTG	CGCGCGGTGC
751	CTCGTGTCT	TGTGGTTTAA	CGCCTATCC	CGCGAAGTCT	TATTGGGCGA
801	TGTCGTTGCA	TGGCATTTGG	GTGCCGCGCT	CGGTACCGTC	GCCGTTATCG
851	TCCGCCAAGA	GTTTGTCTCT	GTCATTATGG	GCGGATTATT	TGTCGTAGAA
901	GCCGTATCCG	TTATGCTTCA	GGTTGGCTGG	TATAAGAAAA	CCAAAAACG
951	CATCTTCTTG	ATGGCGCCCA	TCCATCACCA	CTACGAACAA	AAAGGCTGGA
1001	AAGAAACCCA	AGTCGTCGTC	CGCTTTTGG	TATTACCATT	CGTCTTGGTG
1051	TTGATCCGTT	TGAGTACCCT	CAAAATCCGC	TGAACCTATG	CCGTCCTGAA
1101	ATCTTTTCAGA	CGGCATTTGA	ACGCGCAATA	A	

1 MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51 KCGQAVRTDG PQTHLVKNGT PTMGGSLLT AITVSTLLWG NWANPYIWIL
101 LGVLLATGAL GFYDDWRKV YKDPNGVSAK FKMWQSSVA VIAGLALFYL
151 AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLT DGLDGL
201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQ PYVAGANEVA IFCTAMCGAC
251 LGFLFVNAPY AQVFMGDVGA LALGAALGT AVIVRQEFVL VIMGGLFVVE
301 AVSVMLHVWG YKTKKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
351 LIGLSTLKIR GNYAVRTPFR RHLNAQ*

m088.ppt

1	MFLWLAHFSX	WLTGLNXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX
51	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX
101	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX
151	XXXXXXXXXX	XXXXXXXXXX	XGVVGFVLVS	YLTIVGTSNA	VNLTDLGLDGL
201	ATFPVVLVAA	GLAIFAYASG	HSQFAQYLQL	PYVAGANEVV	IFCTAMCGAC
251	LGFLVFNAPY	AQVFMGDVGA	LALGAALGTV	AVIVRQEFVL	VIMGGFLVVE
301	AVSVMLQVGW	YKTKTKRIFL	MAPIHHHYEQ	KGWKETQVVV	RFWIITVLV
351	LIGLSTLKIR	XTYAVXTSFR	RHLNAQ*		

ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng) from *N. gonorrhoeae*:

m088/g088

m088.pep
 GVVGFVLVLSYLTIVGTSNAVNLTGDLGLA
 |||||
 g088
 IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTGDLGLA
 150 160 170 180 190 200

282

```

      40      50      60      70      80      90
m088.pep  TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g088      AAFPVVLVAAGLAIFAYVSGHYQFSQYLQLPYVAGANEVAIFCTAMCGACLGFLWFNAYPA
          210      220      230      240      250      260

      100     110     120     130     140     150
m088.pep  QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGFLVVEAVSVMLQVGWYKTKKRIFLM
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g088      QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGFLVVEAVSVMLHVGWYKTKKRIFLT
          270     280     290     300     310     320

      160     170     180     190     200
m088.pep  APIHHHYEQKGWKETQVVVRFWIIITIVLVLIGLSTLTKIRXTYAVXTSFRRLNAQX
          |||||: : |||||:|||||:|||||:|||||:|||||:|||||:|||||
g088      APIHHHYQLRCWKETQVVVRFWIIITIVVVLIGLSTLTKIRGNVAVRTPFRRLNAQX
          330     340     350     360     370

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 287>:

```

a088.seq
1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51  TTTTCAATAC ACCACATTCC GCGCCGTCAT GCGGCGTGTG ACCGCCTTGG
101 CGTTTTCCCT GATGTTCTGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
151 AAATGCGGGC AGGCAGTGCG TACCGACGGT CCGCAAACCC ACCTCGTCAA
201 AAACGGCACG CCGACGATGG GCGGTTTCGCT GATTCTGACC GCCATTACCG
251 TGTCCACCCT GTTGTGGGGC AACTGGGCAA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAT TGCTCGCCAC GGGCGCACTC GGTTTTTACG ACGACTGGCG
351 CAAAGTCGTC TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAATGG
401 TGTGGCAGTC AAGCGTTGCC ATTATCGCCG GTTTGGCATT GTTTTACCTT
451 GCGGCCCAAT CCGCCAACAA TATTTTGATT GTCCCGTTCT TCAAACAAAT
501 CGCCCTGCCG CTGGGCGTGG TCGGCTTTTT GGTGTTGTCT TACCTGACCA
551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
601 GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
651 TGCCAGCGGC CACTCACAAT TTGCCAATA CCTGCAATTA CCTTACGTTG
701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTGGTGCA TTGGCATTGG GTGCCGCGCT CCGTACCGTC GCCGTCATCG
851 TCCGCCAAGA GTTTGTCTCT GTCATTATGG GCGGATTATT TGTCGTAGAA
901 GCCGTATCCG TTATGCTTCA GGTCGGCTGG TATAAGAAAA CCAAAAAACG
951 CATCTTCTG ATGGCGCCCA TCCATACCA CTACGAACAA AAAGGCTGGA
1001 AAGAAACCCA AGTCGTCGTC CGCTTTTGA TTATTACCAT CGTCTTGGTG
1051 TTGATCGGTT TGAGTACCCT CAAAATCCGC TGAACCTATG CCGTCTGAAC
1101 ACCTTTCAGA CGGCATTGTA ACGCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 288; ORF 088.a>:

```

a088.pep
1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
101 LGVLLATGAL GFYDDWRKV VYKDPNGVSAK FKMVWQSSVA I IAGLALFY L
151 AANSANNILI VPFFKQIALP LGVVGFVLVS YLTIVGTSNA VNLTDGLDGL
201 ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGFLVVE
301 AVSVMLQVGW YKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
351 LIGLSTLKIR *TYAV*TPFR RHLNAQ*

```

m088/a088 99.5% identity over a 205 aa overlap

```

      150      160      170      180      190      200
m088.pep  XXXXXXXXXXXXXXXXXXXXXXXXXXXXGVVGFVLVLSYLTIVGTSNAVNLT DGLDGLA
          |||||
a088      IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFVLVLSYLTIVGTSNAVNLT DGLDGLA

```

283

	150	160	170	180	190	200
	210	220	230	240	250	260
m088.pep	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA					
a088	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260
	270	280	290	300	310	320
m088.pep	QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGFLVVEAVSVMLQVGWYKTKKRIFLM					
a088	QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGFLVVEAVSVMLQVGWYKTKKRIFLM					
	270	280	290	300	310	320
	330	340	350	360	370	
m088.pep	APIHHHHEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRHLNAQX					
a088	APIHHHHEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTPFRRHLNAQX					
	330	340	350	360	370	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 289>:

```

g089.seq
1  ATGCCGCCCA AAATCACGAA GAGCGGGTTT TGCAAACCGG CAATCGCGGC
51  GGCGGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATG AATACCACGC
101 CGTTTTTCTC GCCGATTTT TCCACACGGT GCGGCAAGCC TTGGAAGGTT
151 TTGACGTGTT CCAGCAATGC TTCGCGCGGC AAACCGACGG CCTCGCACAA
201 AGCCACGGCA GCCATAACGT TGGCGGCGTT GTGCAAACCT TGCAGCGGGA
251 TGTCTTGCGT AGAAATCAAA TCTTCATTGC CTTGTTTTAA ACAGCCCGTC
301 CCGCGTTCCA ACCAAAAATC GGCTTCGTGT TCCAAGGAAA ACCGTTTTCAC
351 TTCACGCCCT GCGCGTTTCA TGGCGCGGCA GAACACGTCG TCCGCATTAG
401 AAACCTGCAC TCCATCGCCA CGGAAAATCT CGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

```

g089.pep
1  MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGKPWKV
51  LTCSSNASRG KPTASHKATA AITLAALCKP CSGMSCVEIK SSLPCFKQPV
101 PRSNQKSASC SKENRFTSRP ARFMARQNTS SAFKTCTPSP RKISALVCA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 291>:

```

m089.seq
1  ATGCCGCCCA AAATCACKAw GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA AACACCACGC
101 CGTTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC TTGGAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG CCTCACACAA
201 AGCCACKGCA GCCATGACGT TAGCGGCGTT GTGCAKACCT TGCAACGGwA
251 TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG GCGGCCTGTC
301 TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA ACCATTTTAC
351 CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCATAG

```

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

```

m089.pep
1  MPPKITXSGF CKPAIAAAVA PTFVPLLSSI NTTPFFSPIF STRCGRPWKV
51  LTCSSNASRD KPMASHKATA AMTLAALCXP CNMGSCVTIK SSLPCFRRPV
101 SRSNQKSASC SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from *N. gonorrhoeae*:

m089/g089

284

	10	20	30	40	50	60
m089 . pep	MPPKITXSGFCKPAIAAAVAPT FVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD					
	: :					
g089	MPPKITXSGFCKPAIAAAVAPT FVPLLSSMNTTPFFSPIFSTRCGKVPWKVLTCSSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089 . pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP					
	: : : : :					
g089	KPTASHKATAAITLAALCKPCSGMSCVEIKSSLPCFKQPVPRSNQKSASCSKENRFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089 . pep	ARFIARQNASSAFKTCTPSRKILALVCAX					
	: :					
g089	ARFMARQNTSSAFKTCTPSRKISALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 293>:

a089 . seq

```

1   ATGCCGCTA AAATCAGAA GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GGCGGTTCG CCGACGTTTC TGCCTTTGCT GTCGTCGATG AACACCACGC
101 CATTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC TTGAAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGGC AAACCGACGG CTTCGCACAA
201 GGCAACGGCA GCCATCACGT TAGTGGCGTT GTGCAAGCCT TGCAGCGGAA
251 TATCTTGCGT GGCAATCAAA TCTTCATTGC CTTGTTTCAG GCGACCTGTC
301 TCACGTTCCA ACCAAAAATC GGCTTCGTAT TCCAACGAAA ACCATTTCAC
351 CTCGCGCCCG GCGCGCTTCA TCGCAGGACA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC ACCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>:

a089 . pep

```

1   MPPKITKSGF CKPAIAAAVA PTEFVPLLSSM NTTPFFSPIF STRCGRP*KV
51  LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV
101 SRSNQKSASY SNEHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

m089/a089 91.9% identity over a 149 aa overlap

	10	20	30	40	50	60
m089 . pep	MPPKITXSGFCKPAIAAAVAPT FVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD					
	: :					
a089	MPPKITXSGFCKPAIAAAVAPT FVPLLSSMNTTPFFSPIFSTRCGRPXKVLTCSSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089 . pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP					
	: : : : : :					
a089	KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPVSRSNQKSASYSNENHFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089 . pep	ARFIARQNASSAFKTCTPSRKILALVCAX					
a089	ARFIARQNASSAFKTCTPSRKILALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 295>:

g090 . seq

```

1   ATGCGCGTAG TCGAGCAAT CGTCGTAGCG GTCGAGATGG TCTTCGAAA

```

g090.pep
1 MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
51 LQFCLQDGRT DIARNDGIQP ALDAEIAQQA GYRGFAVAAG NRNHLVAAAV
101 HNVROOLDVA XHAXRRFA*

```
m090.seq
1  ATGCGCATAG  TCGAGCAAGT  CGTCGTAGCG  GTCGAGATGG  TCTTCGAAA
51  TGTTCAGCAC  CGTCGCCGCA  GTCGGACGCA  GGCTTTCGGT  GTTTCAGT
101  TGGAAAGCTG  AAAGCTCCAA  CAGCCACACG  TCCGCCTTTT  TCTTCGCG
151  CTGCCATTCC  GCCTCCAAAA  CCGGCGTGCC  GATATTGCCC  GCGATAACGG
201  TATCCAGCCC  GCACTTGATA  CAGAGATAGC  CGACCAGGCT  CGTTACCGTG
251  GTTTTGCCGT  TGCTGCCGGT  AATCGCAATT  ACCTTGTCGT  CCCGGCGGTT
301  CACAATGTCC  GCCAGCAATT  CGATTGTCGC  CAACACGCGT  .CCGCCGTTT
351  TGCTTGA
```

```
m090.pep
  1  MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
 51  LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNYLVVPAV
101  HNVROQFDVA OHAXRRFA*
```

ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng) from *N. gonorrhoeae*:

m090/g090

	10	20	30	40	50	60
m090.pep	MRIVEQVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRRA					
	: :	:	:	:	:	: :
g090	MRVVEQIVVAVEMVFGNVHRRRSRAQAFGVFQLEAGKLPHPHVRLFAFALQFCLQDGRT					
	10	20	30	40	50	60

	70	80	90	100	110	119
m090.pep	DIARDNGIQPALDTEIADQARYRGFAVAAGNRNYLVVPVHNVRQQFDVAQHAXRRRFAX					
	: :	: :	:	: :	: :	:
g090	DIARNDGIQPALDAEIAHQAGYRGFAVAAGNRNHLVAAVHNVRQQLDVAXHAXRRRFAX					
	70	80	90	100	110	

a090.seq

```
1 ATGCGCGTAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGA
51 TGTTACAGCAC TGTCGCCGCA GTCGGGCGCA GGCCTTCGGT GTTTCACAG
101 TGGAAACTGG AAAGCTCCAA CACCACACG TCCGCCTTTT TGCTTCGCG
151 CTGCAATTCC GCCTCCAAAA CCGGCGCGCC GATATTGCCC GCGATAACGG
201 TATCCAGCCC ACACTTGATG CAGAGATAGC CGACCAGGCT CGTTACCGTG
251 GTTTTGCCGT TGCTGCCGGT AATCGCAATC ACCTTGTCGC CGCGGCGGTT
301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT C.CGCGGTTT
351 CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 300; ORF 090.a>:

a090.pep

```

1 MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA
51 LQFRLQNRRA DIARDNGIQP TLDAEIA DQA RYRGFAVAAG NRNHLVAAAV
101 HNVRRQQFDVA QHAXRRFA*

```

m09/a090 91.5% identity over a 117 aa overlap

```

              10      20      30      40      50      60
m090.pep    MRIVEQVVVA VEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA
              ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a090         MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQHPHVRLFAFALQFRLQNRRA
              10      20      30      40      50      60

              70      80      90     100     110     119
m090.pep    DIARDNGIQPALDTEIADQARYRGFAVAAGNRNLYLVVPAVHNVRQQFDVAQHAXRRFAX
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a090         DIARDNGIQPTLDAEIA DQARYRGFAVAAGNRNHLVAAAVHNVRQQFDVAQHAXRRFAX
              70      80      90     100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae*

g090-1.seq This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence < ORF 090-1.ng>:

g090-1.pep (not shown)

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 2>:

m090-1.seq

```

1 ATGACGGCGT TTGCATTCA GACGGCATCA CAAAGCCTTA AACGCTTCGA
51 TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG
101 CAGGCGGGGC TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCATATGC
151 CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC
201 AGCCGTCCAA ATCGCGGGCG ATTTGCGGCG CATCGACACC AATCAAGAAC
251 ACGCCTTTTG CCTTGCCCTAC CAGTGCATCG CGCAGGGGCG TGAAGTCCTG
301 CCCTTTACCC ATGCCGCCCA AAATCACGAA GAGCGGATTT TGCAAACCGG
351 CAATCGCGGC GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA
401 AACACCACGC CGTTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC
451 TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG
501 CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGCGGTT GTGCAGACCT
551 TGCAACGGAA TGTCTTGCGT GACAATCAA TCTTCATTGC CTTGTTTCAG
601 GCGGCCTGTC TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA
651 ACCATTTTAC CTCGCGCCCG GCACGTTCA TCGCGCGGCA GAACGCATCG
701 TCCGCATTCA AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT
751 ATGGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGA
801 TGTTACGAC CCGCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
851 TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCCTCGCG
901 CTGCCATTCC GCCTCCAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
951 TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAAGCT CGTTACCGTG
1001 GTTTTGCCGT TGCTGCCGCT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
1051 CACAATGTCC GCCAGCAATT CGATGTGCGC CAACACGCGT CCGCCGTTTT
1101 GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA
1151 GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA
1201 AAACACCAAC CCGTCAAACA TCTTACCGAT TTGCGACACG CGTTCGGCT
1251 TCAGCTCCGC ATCATACGCA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG
1301 GCAATCATGG AAATACCCGT ACCGCCGAGT CCGGCGACGA GGATTTTTTT
1351 GTTTTGAAAA GTCATTTTGG TTTGTCCTAA

```

This corresponds to the amino acid sequence <SEQ ID 3; ORF 090-1>:

m090-1.pep

```

1 MTAFAFQTAS QSLKRFDKHF RTVRVAFEHI KARAGGAEQH NIACFGLGIC
51 RLNGFSQSGA VGHQAAAVQ IAADLRRIDT NQEHAFCLAY QCIAQGREV
101 PFTHAAQNHE ERILQTGNRG GSRADIRAF VVDKHHAVFL ADFHFAVRQA
151 LEGFDVFEQC FARQTDGLTQ SHGSHDVSGV VQTLQRNVLR DNQIFIALEQ
201 AACLAFQPEI SFVFQRKPFY LAPGTLHRAA ERIVRIQNLH AVATENLGF
251 MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
301 LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNLYLVVPAV
351 HNVRRQQFDVA QHASAVLLER LNIRLPLADA GTESQNIETV VQRIFQTARV
401 KHQPVKHLTD LRHAFRLQLR IIRSNLRAVF AQVGNHGNTR TAESGDEDEFF
451 VLKSHFGLS*

```

This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 305>:

This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng) from *N. gonorrhoeae*:

```

                                10      20      30      40      50      60
m091.pep  MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPLPKPLSDGIASCSITRLQALVIVAA
          |||||:| |||||:| ||||| ||||| |||||
g091      MEIPVPPSPATRIFLFESHFGFVLKQIILGRRRPPLPKPLSDGIASRLITRLQALVIVAA
          10      20      30      40      50      60

                                70      80      90      100
m091.pep  VLVSVLTSLAKPFLCKGAVLAHAASFGIHHQIVLGLGYPLR
          |||||:| : |||||:| |:| ||||| ||
g091      VLVSVLTSLAKPLLSEKVLHAHAASIVIHQAQIVLGLGIPLEX
          70      80      90      100

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 307>:

a091.seq	1	51	101	151	201	251
	ATGGAATAC	CCGTGCCGCC	AAGTCCGGCG	ACGAGGATTT	TTTGTTTTG	
	GAAATCATTT	TGGTTTGTC	TAAAACAAAT	CATATTGAGC	AGGGGATGTC	
	TGATCCTGCT	CAAGCCGCTT	TCAGACGGCA	TCGCGAGCTG	TCGAATAACC	
	CGCTTTTCAG	CGTTGGTCAT	TGTCGCAGCT	GTCTTGGTAT	CCGTTTTGAC	
	AAGCCTTGCC	AAGCCATTCT	TGTGCAAGGG	CGCGGTCTTG	GCGCACGCCG	
	CGTCTTTTCG	CATACATCAC	GCCCAAATTG	TTTTGGGC		

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>:

a091.pep

```
1 MEIPVPPSPA TRIFLWKSF WFVLKQIILS RGCLILLKPL SDGIASCSIT
51 RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG
```

m091/a091 93.8% identity over a 96 aa overlap

	10	20	30	40	50	60
m091.pep	MEIPVPPSPATRIFLFEKSF	WFVLKQIILSRRCPPLPKPLSDGIASCSIT	RLQALVIVAA			
a091	MEIPVPPSPATRIFLFEKSF	WFVLKQIILSRGCLILLKPLSDGIASCSIT	RFQALVIVAA			
	10	20	30	40	50	60
	70	80	90	100		
m091.pep	VLVSVLTSLAKPFLCKGAVLAHAASFGIHH	AQIVLGLGYPLR				
a091	VLVSVLTSLAKPFLCKGAVLAHAASFGIHH	AQIVLG				
	70	80	90			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 309>:

g092.seq

```
1 ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
51 AAACGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCAACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTTGTCCGT
151 ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGGTTCCG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
301 AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAAGA AAGAAaacc
351 CGAAGTgtgc gcTGCCTTGG AGCGGCAAAT TCCCGTTATT CCGCGCGCCT
401 TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACGgcacgc cattgcccgt
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAATC AACGCCGCAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTGTTGACAG CGAACACGTC CGCGCGATTG TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGG CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCGCAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGc gtggcGCTg
1001 aagtcGgCGC ATcggttgAA GCGAtcCAA AaggCTTGCT CGGCTTTGAA
1051 GCGCTCGGCC GCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
1101 cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCCC GTCGAAATGG
1151 CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGGAAAA acgtTTGGTG
1201 CtcgCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTAACAATA CCGTTGatgC GCTGGTACTG ACCGAAGTTT
1301 AtgccgccgG CGAAGAGCCG GTTGCCGCCG CCGactcCCG CGCCTTGGCG
1351 CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTTACT GCGAAAatgt
1401 cgccgACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT Ggcgatgttg
1451 tgttgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
1501 gaattgtcga AACAGAttg A
```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

g092.pep

```
1 MFFISIRYIF VRKLWCANGQ TFKITPLRTK NQPERNIMMK NRVSNIHFVG
51 IGGVMSGIA EVLHNLGFKV SGSDQARNAA TEHLSSLGIQ VYPGHAEHV
101 NGADVIVAST AVKKENPEVV AALERQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
```

```

251 AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKTIVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
451 RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
501 ELSKQI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 311>:

```

m092.seq
1  ATGTTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51  AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTTA CCAACATCCA TTTTGTCTGGT
151 ATCGGCGGCG TCGGCATGAG CGGCATCGCC GAAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGGTTCCG ATCagGCGCG AAATGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
301 AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551 GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGCGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCCT TGCCGCCGCA CGCGCGCGCT ATCTGAAAAA ACGTTTGGTA
1201 CTCGCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTTGGG CAAACTCGAG CCGATTACTT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGAACGT TTTGCAGGAC GGCGACATCG
1451 TGTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCCGC CGCGTGCTG
1501 GCATTGTCGA AACAGATTG A

```

This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:

```

m092.pep
1  MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFVG
51  IGGVMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHAEHV
101 NGADVVTST AVKKNPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKTIVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMAG SINRVPAALL
501 ALSKQI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from *N. gonorrhoeae*:

```

m092/g092
10      20      30      40      50      60
m092.pep  MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVMSGIA

```



```

|||||
g092      MFFISIRYIFVRKLWCANGQTFKITPLRTKNQPERNIMMKNRVSNIHFGVIGGVGMSGIA
              10      20      30      40      50      60

              70      80      90      100     110     120
m092.pep   EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNAGDVVVTSTAVKKENPEVV
              |||||
g092      EVLHNLGFKVSGSDQARNAATEHLSSLGIQVYPGHTAEHVNAGDVVVASTAVKKENPEVV
              70      80      90      100     110     120

              130     140     150     160     170     180
m092.pep   AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
              |||||
g092      AALERQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
              130     140     150     160     170     180

              190     200     210     220     230     240
m092.pep   NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
              |||||
g092      NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
              190     200     210     220     230     240

              250     260     270     280     290     300
m092.pep   FIHRMPFYGKAFLCIDSEHVRILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
              |||||
g092      FIHRMPFYGKAFLCVDSEHVRILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV
              250     260     270     280     290     300

              310     320     330     340     350     360
m092.pep   QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG
              |||||
g092      QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG
              310     320     330     340     350     360

              370     380     390     400     410     420
m092.pep   DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLAFQPHRYTRTRDLFEDFTK
              |||||
g092      DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYPEKRLVLAFQPHRYTRTRDLFEDFTK
              370     380     390     400     410     420

              430     440     450     460     470     480
m092.pep   VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD
              |||||
g092      VLNTVDALVLTEVYAAGEEPVAAADSRALARAIRVLGKLEPIYCENVADLPQMLMNVLQD
              430     440     450     460     470     480

              490     500
m092.pep   GDIVLNMGAGSINRVPAALLALSQIX
              ||:|||||
g092      GDVVLNMGAGSINRVPSALLELSQIX
              490     500

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 313>:

a092.seq

```

1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51  AACGGTCAG  CCCTTTAAAA TAACGCCTTT  ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTGA CCAACATCCA TTTTGTCGGT
151 ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
201 TTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT

```

```

301 AACGGTGC GG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGT GTTGAGAAGC
701 TGCATCAGGC GTTATATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGCAGACAC AACGTGCTGA ACGCATTTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGTGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 TGGAACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCTT TTCCGCCGCA CGCGGCGCGT ATCCGAAAAA ACGTTTGGTA
1201 CTCGCCTTCC AGCCGCACCG CTATACCGCG ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCGCG TGAAGAGCCG ATTGCCGCCG CTGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTTGGG CAAACTCGAG CCGATTTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGTAACGT TTTGCAGGAC GGCGACATCG
1451 TGTGTAATAT GGGTGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501 GAATTGTCGA AACAGATTTG A

```

This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>:

a092.pep

```

1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFGV
51 IGGVGMMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKNPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKTIVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVRRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLSAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
501 ELSKQI*

```

m092/a092 99.4% identity over a 506 aa overlap

	10	20	30	40	50	60
m092.pep	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFGVIGGVGMMSGIA					
a092	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFGVIGGVGMMSGIA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m092.pep	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKNPEVV					
a092	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKNPEVV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m092.pep	AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL					
a092	AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m092.pep	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
a092	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					

292

	190	200	210	220	230	240
	250	260	270	280	290	300
m092.pep	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTIVHV					
a092	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTIVHV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m092.pep	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG					
a092	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG					
	310	320	330	340	350	360
	370	380	390	400	410	420
m092.pep	DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLAQPHRYTRTRDLFEDFTK					
a092	DIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKRLVLAQPHRYTRTRDLFEDFTK					
	370	380	390	400	410	420
	430	440	450	460	470	480
m092.pep	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
a092	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
	430	440	450	460	470	480
	490	500				
m092.pep	GDIVLNMGAGSINRVPAALLALSKQIX					
a092	GDIVLNMGAGSINRVPAALLELSKQIX					
	490	500				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 315>:

g093.seq

```

1  atGCAGAAAtt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
51  ACGAGAAatc tcgcTGGACA GCgGTACCGC CATTTTGAAC GCCTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTTATCC
151 GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAATG TTGGGCATTG
251 CCTATACCGG CAGCGGTGTC GCCGCTCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTGGCA GGCATTGGGA TTACCCGTTT CCGAGTTCGC
351 CGTACTGTAC GATGATACCG ATTTTCGATG CGTCGAAGAA AAATTGGGTC
401 TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggtta
451 aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtacgaag aatTGAaaCA
501 CCTTcagggg cgaAatcatt gccgAacgTT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATA ACATCATCCC
601 CGCAACCGAG TTTTACGAct acgaagccaa GtacaaCCGA GACGAcacca
651 tttatCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGGAACTGG CGGTTGCGG CGCACAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACCG CAAACTCTAT CTGTTGGAAA
801 TCAACACCCT GCCCGGTATG ACCGGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

g093.pep

```

1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKERGFQTA FNILHGTIGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLY DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
151 KVKEKGR LKS VYEELKHLQG RNHCRTFYRR RRIFLPRPER QRAARHTHHP
201 RNRVLRLRSQ VQPRRHLSM SFGRFDRSRR KPDARTGGSR RTGNRCGLRL
251 ARRFQRYRR QTLVGNQHP ARYDRP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 317>:

m093.seq

```

1  ATGCAGAATT TTGGCAAAGT GCCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CTAAAGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCrAA GACGGGGCGG TTCAGGGTGC ATTGGAAGTGTG TTGGGCATTTC
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGTGCAAAC TGATTTGGCA GGCATTGGGA TTGCCCCTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTTCGATG CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501 CCTTCAGGG. CGAAATCATT GCCGAACGTT TTATCGGCGG CCGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATA ACATCATTCC
601 CGCAACCGAG TTTTACGACT ACGAAGCCAA GTACAACCGC GACGACACCA
651 TTTATCAATG TCCTTCGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTTCGGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAG
801 TCAACACCCT GCCCGGTATG ACGAGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 318; ORF 093>:

```

m093.pep
1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGX DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGLSSVGVV
151 KVKGKGR LKS VYEELKHLQX RNHCRTFYRR RRIFLPRPER QRAARHTHHS
201 RNRVLR LRSQ VQPRRHLSM SFGFRDRSRR KPDARTGGSR RAGNRCGRLR
251 ARRFQRYRR QTLVGNQHP ARYDEP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from *N. gonorrhoeae*:

```

m093/g093
      10      20      30      40      50      60
m093.pep  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPL SELKAQGFQTA
          |||
g093      MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPL SELKERGFQTA
          |||
      10      20      30      40      50      60

      70      80      90     100     110     120
m093.pep  FNILHGTYGX DGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
          |||
g093      FNILHGTYGEDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLV
          |||
      70      80      90     100     110     120

      130     140     150     160     170     180
m093.pep  DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKGKGR LKSVYEELKHLQXRNHCRTFYRR
          |||
g093      DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVKVKEKGR LKSVYEELKHLQGRNHCRTFYRR
          |||
      130     140     150     160     170     180

      190     200     210     220     230     240
m093.pep  RRIFLPRPERQRAARHTHHSRNRVLR LRSQVQPRRHLSMSFGRFDRSRRKPDARTGGSR
          |||
g093      RRIFLPRPERQRAARHTHHPRNRVLR LRSQVQPRRHLSMSFGRFDRSRRKPDARTGGSR
          |||
      190     200     210     220     230     240

      250     260     270
m093.pep  RAGNRCGRLRARRFPQRYRRQTLVGNQHPARYDEPX
          |||
g093      RTGNRCGRLRARRFPQRYRRQTLVGNQHPARYDRPX
          |||
      250     260     270

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 319>:

```
a093.seq
1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CCAAGGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAACG TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTTCGATG CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501 CTTTCAGGG. CGAAATCATT GCCGAACGGT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCTG TGTTGAACGG CAAAGGCCTG CCCGGCATAC ACATCATCCC
601 CGCGACCGAG TTTTATGACT ACGAAGCCAA GTACAACCGC AACGACACCA
651 TTTATCAATG TCCTTCGGAA GATCTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTTCGGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAG
801 TCAACACCCT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>:

```
a093.pep
1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGE DGA VQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AEGSSVGVV
151 KVKGKGR LKS VYEELKHFXQ RNHCRTVYRR RRIFLPCVER QRPARTHHF
201 RDRV L* LRSQ VQQRHHL SM SFGRSDRSRR KPDARTGGS RAGNRCGR LR
251 ARRF PQR YRR QTL SVGNQHP ARYDRP*
```

m093/a093 95.7% identity over a 276 aa overlap

	10	20	30	40	50	60
m093.pep	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPL SELKAQGFQTA					
a093	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPL SELKAQGFQTA					
	10	20	30	40	50	60
m093.pep	FNILHGTYGXDGA VQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
a093	FNILHGTYGEDGA VQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
	70	80	90	100	110	120
m093.pep	FNILHGTYGXDGA VQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
a093	FNILHGTYGEDGA VQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
	130	140	150	160	170	180
m093.pep	DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVVKVGKGR LKSVYEELKHLQXRNHCRTFYRR					
a093	DDTDFDAVEEKLGLPMFVKPAAEGSSVGVVVKVGKGR LKSVYEELKHFXQRNHCRTVYRR					
	130	140	150	160	170	180
m093.pep	RRIFLPRPERQRAARH THSRNRLRLRSQVQPRRHLSMSFGRFDRSRRKPDARTGGS R					
a093	RRIFLPCVERQRPART HHPRDVLXLR SQVQQRHLSMSFGRSDRSRRKPDARTGGS R					
	190	200	210	220	230	240
m093.pep	RRIFLPRPERQRAARH THSRNRLRLRSQVQPRRHLSMSFGRFDRSRRKPDARTGGS R					
a093	RRIFLPCVERQRPART HHPRDVLXLR SQVQQRHLSMSFGRSDRSRRKPDARTGGS R					
	190	200	210	220	230	240
m093.pep	RAGNRCGR LRARRFPQRYRRQTL SVGNQH PARYDEPX					
a093	RAGNRCGR LRARRFPQRYRRQTL SVGNQH PARYDRPX					
	250	260	270			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 321>:

```
g094.seq
  1 ATGTATTTCG CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
 51 GCCGCCGATA ACGAAAGTGG GGTTCGAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTAccggc aatggcgatg
151 cCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
201 GGAATTTGCG CGCTCCAACG CAgcgacaAC TTCGGgattT TCTTTCTTGA
251 CGGCGGTAGA GGCAACGACG ACATccgcAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:

```
g094.pep
  1 MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
 51 PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTCSAV
101 WPG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 323>:

```
m094.seq
  1 ATGTATTTCG CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
 51 GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGCG AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGAATTTGCG TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTTTGA
251 CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCGGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:

```
m094.pep
  1 MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
 51 PSRKRISSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101 WPG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 094 shows 95.1% identity over a 103 aa overlap with a predicted ORF (ORF 094.ng)

from *N. gonorrhoeae*:

```
m094/g094

      10      20      30      40      50      60
m094.pep  MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAN
          |||
g094      MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAS
          |||

      70      80      90     100
m094.pep  IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
          |:|
g094      IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTCSAVWPGX
          |:|

      70      80      90     100
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 325>:

```
a094.seq
  1 ATGTATTTCG CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
 51 GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGCG AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGAATTTGCG TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTTTGA
251 CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

a094.pep

```

1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101 WPG*

```

m094/a094 100.0% identity over a 103 aa overlap

```

              10      20      30      40      50      60
m094.pep      MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
               |||
a094           MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
               |||
              10      20      30      40      50      60

              70      80      90     100
m094.pep      IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
               |||
a094           IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
               |||
              70      80      90     100

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 327>:

```

g095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AACACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA
301 GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTGCGC TTGCCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>:

```

g095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRQCQRK
101 EASDRRLRQR CIRLCPSGRW CLRR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 329>:

```

m095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTGCGC TTGCCCCAAG
351 CGGGCGTTAG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>:

```

m095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRQCQRK
101 DASDRRLRQR CIRLCPSGRX CLRR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae*:

m095/g095

```

              10      20      30      40      50      60
m095.pep      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
               |||

```

297

```

g095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60
           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
g095      HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKEASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
g095      CLRRX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 331>:

```

a095.seq
  1  ATGTCCTTTC  ATTTGAACAT  GGACGGTGAA  TTTCATTTGC  GCGCCGACGT
 51  TTTCGATGTC  GGTGGCGTAG  ATGTCGGCGG  TATCGTCCAA  ACCGTAAGTA
101  GCATAAGGTT  TGCTCACTTT  GGGCAAATC  GCGCGGACGT  GTTCGCTGTC
151  AATACACAAA  AAGGCTTTGC  CGTAGAAGGG  CATACGGTGG  ATGAAATCGA
201  TAAACGCCTG  ATGCAGCTTC  TCAACACTGT  GCCCGTAGGT  ATCCATATGG
251  TCTTCGTCGA  TATTGGTAAC  GACGGACATA  ATCGGTGTCA  GTGCAGAAAG
301  GATGCATCCG  ACCGTCGGCT  TCGGCAACGA  TGTATTGCC  TTTGCCCAAG
351  CGGGCGTTGG  TGCCTGCGGC  GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 332; ORF 095.a>:

```

a095.pep
  1  MSFHLNMDGE  FHLRADVFDV  GGVDVGGIVQ  TVSSIRFAHF  GQNRADVFAV
 51  NTQKGFAVEG  HTVDEIDKRL  MQLLNTVPVG  IHMVFDIGN  DGHNRCQCRK
101  DASDRRLRQR  CIRLCPSGRW  CLRR*

```

m095/a095 96.0% identity in 124 aa overlap

```

           10      20      30      40      50      60
m095.pep  MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           |||||
a095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60
           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
a095      HTVDEIDKRLMQLLNTVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
a095      CLRRX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 333>:

```

g096.seq
  1  ATGGCCGGTC  ATACCGGGCA  GGGTGTGAT  TTCCAACAGA  TAGAGTTTGC
 51  CGTCGGTATC  TTTGAGGAAA  TCGACGCGCA  CGCAGCCTTC  CGCACCATT
101  GCCTGTGCGC  CGCGAACCGC  CAGTTCGCGC  ATCAGGCTTT  CTTCCGCTTC
151  GGTCAAATCT  TCCGAAGGAC  ATTGAtaat  ggtgTCGTCT  CGGttgtaCt

```



```

201  tggcttcgta gTCGTAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251  AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301  AcgtTcgga atgaTTtcgc ccctgAAGGT GttTCAattc ttcgtaAACG
351  CTTTTCAGAc ggccttTTTC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>:

g096.pep

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAFF RTDCLCAANR QFAHQAFFGF
51  GQIFRRTLIN GVVSVVLGFV VVKLGCGDDV YAGQPFQAVQD GAGIFAAADK
101 TFGNDFAPEG VSILRKRFSD GLFL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 335>:

m096.seq

```

1  ATGGCTCGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACC GCAGTTCGCGC ATCAGGCTTT CTTGCGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTCG CGGTTGTACT
201 TGGCTTCGTA GTCGTAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 ACGTTCGGCA ATGATTTTCG CC.TGAAGGT GTTCAATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>:

m096.pep

```

1  MARHTGQGVDFQQIEFAVGI FEEIDAHAFF RTDCLRAANR QFAHQAFFGF
51  GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFQAVQD GAGIFAAADK
101 TFGNDFAXEG VSILRKRFSD GLFL*

```

m096/g096 96.0% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGI	FEEIDAHAFFRTDCLRAANR	QFAHQAFFGFGQIFRRTLIN			
g096	MAGHTGQGVDFQQIEFAVGI	FEEIDAHAFFRTDCLCAANR	QFAHQAFFGFGQIFRRTLIN			
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDV	YAGQPFQAVQDGAGIFAAADK	TFGNDFAXEGVSILRKRFSD			
g096	GVVSVVLGFVVVKLGCGDDV	YAGQPFQAVQDGAGIFAAADK	TFGNDFAPEGVSILRKRFSD			
	70	80	90	100	110	120
m096.pep	GLFLX					
g096	GLFLX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 337>:

a096.seq

```

1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACC GCAGTTCGCGC ATCAGGCTTT CTTGCGCTTC
151 GGTCAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTTG CGGTTGTACT
201 TGGCTTCGTA GTCATAAAAC TCGGTCGCGG GGATGATGTG TATGCCGGGC
251 AGGCCTTTGC CGTTCAACAC AGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 CCGTTCGGCA ATGATTTTCG CCT.GAAAGT GTTCAATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>:

a096.pep

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAFF RTDCLRAANR QFAHQAFFGF

```

299

51 GQIFRRTLIN GVVAVVLGFV VIKLGRGDDV YAGQAFVQH RAGIFAAADK
 101 PFGNDFAXES VSILRKRFSD GLFL*

m096/a096 92.7% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDCLRAANRQFAHQAFFGFGQIFRRTLIN				
a096	MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDCLRAANRQFAHQAFFGFGQIFRRTLIN				
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPFVQDGAGIFAAADKTFGND	FAXEGVSILRKRFSD				
a096	GVVAVVLGFVVIKLGCGNDVYAGQAFVQHRAGIFAAADKPF	GNDFAXESVSILRKRFSD				
	70	80	90	100	110	120
m096.pep	GLFLX					
a096	GLFLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 339>:

g097.seq

1	ATGGATATTT	CAAACAAAC	ATTGCTGGAT	AGGGTTTTTA	ACCTGAAGGC
51	AAACGGTACG	ACGGTACGTA	CCGAGTTGAT	GGCGGGTTTG	ACGACCTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAATC	CCCTGATTTT	GGGCGAGACC
151	GGAATGGATA	TGGGGGCGGT	ATTCGTCGCT	ACCTGTATCG	CATCCGCCAT
201	CGGCTGTTTT	GTCATGGGTT	TTATCGGCAA	CTATCCGATT	GCGCTTGCCC
251	CGGGGATGGG	GCTGAATGCC	TATTTACCT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCCT	GGCAGGTGGC	GTTGGGTGCG	GTGTTCAATT	CCGGTCTGAT
351	TTTCATCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCTATGGG	TTTGAAAATG	TCGATTGCCG	CCGGTATCGG	TTTGTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCGGC	TTGGGCGATA	TTCATCAGCC	CAGCGCACTG	TTGGCATTGT
551	TCGTTTGT	CATGGTGGTC	GTATTGGGGT	ATTTCCGCGT	TCAAGGCGCA
601	ATCATCATCA	CCATTCTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAGTTT	CACGGCGTGG	TCGGCGAAGT	ACCGGGCATT	GCGCCGACCT
701	TTATGCAGAT	GGATTTTAAA	GGTCTGTTTA	CCGTCAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TCTTCTTGGT	CGATTTGTTC	GACAGTACCG	GAACGCTGGT
801	CGGCGTATCC	CACCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCCGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CCATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGCGCGG	CGGGCGTATC
951	GGCAGGCGGA	CGGACCGGCC	TGACGGCGGT	TACCGTCGGC	GTATTGATGC
1001	TGGCGTGTCT	GATGTTCTCC	CCATTGGCGA	AAAGTGTTCC	GGTATTTGCC
1051	ACCGCGCCCG	CACTGCTTTA	TGTCGGCACG	CAGATGCTCC	GCAGTGCGAG
1101	GGACATTGAT	TGGGACGATA	TGACTGAAGC	CGCGCCCGCG	TTCTTGACCA
1151	TTGTCTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCCTTCGGC
1201	TTCATCAGCT	ATGCCGTGGT	CAAACTTTTG	TGTCGCCGGA	CTGGGGACGT
1251	GCCGCCTATG	GTATGGGTTG	TTGCCGTATT	GTGGGCATTG	AAATTCTGGT
1301	ATTTGGGCTG	A			

This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:

g097.pep

1	MDISKQTL	LD RVFNLKANGT	TVRTELMAGL	TTFLTMCYIV	IVNPLILGET
51	GMDMGAVFVA	TCIASAIGCF	VMGFIGNYPI	ALAPGMGLNA	YFTFAVVKGM
101	GVPWQVALGA	VFISGLIFIL	FSFFKVREML	VNALPMGLKM	SIAAGIGLFL
151	ALISLKGAGI	IVANPATLVG	LGDIHQPSAL	LALFGFVMV	VLGYFRVQGA
201	IIITILTITV	IASLMGLNEF	HGVVGEVPGI	APTFMQMDFK	GLFTVSMVSV
251	IFVFFLVDLF	DSTGTLVGVS	HRAGLLVDGK	LPRLKRALLA	DSTAIVAGAA
301	LGTSSTPYV	ESAAGVSAGG	RTGLTAVTVG	VLMLACLMS	PLAKSVPVFA
351	TAPALLYVGT	QMLRSARDID	WDDMTEAAPA	FLTIVFMPFT	YSIADGIAFG

300

401 FISYAVVKLL CRRTGDVPPM VWVAVLWAL KFWYLG*The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 341>:

m097.seq

```

1   ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
51  AAACGGTACK ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
151 GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
251 CGGGGATGGG GCTGAATGCC TATTTACCTT TTGCCGTCGT TAAGGGTATG
301 GCGGTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTTCATCT CCGGTCTGAT
351 TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCATTGT
551 TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAATGAATTT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
701 TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCAGTGTG
751 ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
801 CGGCATATCC CACCGTGCCG GGCTGCTGGT GGACGGTAAG CTGCCCCGCC
851 TGAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCCACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTC CGCTTTTGCC
1051 ACCGCGCCCG CCCTGCTTTA TGTCGGCAGC CAGATGCTCC GCAGTGCGAG
1101 GGATATTGAT TGGGACGATA TGACGGAAGC CGCACCTGCG TTCCTGACCA
1151 TTGTTTTTCAT GCCGTTTACT TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCGCA CCAAAGACGT
1251 TCCGCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301 ATTTGGGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

m097.pep

```

1   MDTSKQTL LD GIFKLKANGT TVRTELMAGL TTFLTMCIYIV IVNPXILGET
51  GMDMGAVFVA TCIASAIGCF VMGFVGNYP I ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201 IITILTITV IASLMGLNEF HGIIGEVP SI APTFMQMDFE GLFTVSMVSV
251 IFVFFLVDFL DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMEF PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVPPM VWIVAVLWAL KFWYLG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng) from *N. gonorrhoeae*:

m097/g097

```

          10          20          30          40          50          60
m097.pep  MDTSKQTL LDGIFKLKANGT TVRTELMAGL TTFLTMCIYIV IVNPXILGET GMDMGAVFVA
          || ||||| : |||||
g097      MDISKQTL LDRVFNKANGT TVRTELMAGL TTFLTMCIYIV IVNPLILGET GMDMGAVFVA
          10          20          30          40          50          60

          70          80          90          100         110         120
m097.pep  TCIASAIGCF VMGFVGNYP IALAPGMGLNAY FTFAVVKGM GVPWQVALG AVFISGLIFIL
          ||||| : |||||
g097      TCIASAIGCF VMGFIGNYP IALAPGMGLNAY FTFAVVKGM GVPWQVALG AVFISGLIFIL
          70          80          90          100         110         120

          130         140         150         160         170         180

```

301

m097.pep	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGIIV	ANPATLV	GLGDIHQ	PSAL
g097	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGIIV	ANPATLV	GLGDIHQ	PSAL
		130	140	150	160	170	180		
m097.pep	LALFGF	AMVVVL	GHFRVQ	GAIIIT	ILTITV	IASLMGL	NEFHGI	IGEVPS	IAPTFM
g097	LALFGF	AMVVVL	GHFRVQ	GAIIIT	ILTITV	IASLMGL	NEFHGI	IGEVPS	IAPTFM
	190	200	210	220	230	240			
m097.pep	GLFTVSM	VSIVFV	FFLVDL	FDSTGT	LVGISH	RAGLLV	DGKLPRL	KRALLAD	STAIVAG
g097	GLFTVSM	VSIVFV	FFLVDL	FDSTGT	LVGVSH	RAGLLV	DGKLPRL	KRALLAD	STAIVAG
	250	260	270	280	290	300			
m097.pep	LGTSS	TPYVES	AAGVS	SAGGRT	GTLAVT	VGVLML	ACL	MFSP	LAKSVPA
g097	LGTSS	TPYVES	AAGVS	SAGGRT	GTLAVT	VGVLML	ACL	MFSP	LAKSVPA
	310	320	330	340	350	360			
m097.pep	QMLRS	ARDIDW	DDMTEA	APFLTI	VFMPFT	YSIADG	IAFGFI	SYAVVK	LLCRRTK
g097	QMLRS	ARDIDW	DDMTEA	APFLTI	VFMPFT	YSIADG	IAFGFI	SYAVVK	LLCRRTK
	370	380	390	400	410	420			
m097.pep	VWIVAV	LWALKF	WYLGX						
g097	VWVAV	LWALKF	WYLGX						
	430								

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 343>

a097.seq

1	ATGGACACTT	CAAAACAAAC	ACTGTTGGAC	GGGATTTT	TAAGCTGA
51	AAACGGTACG	ACGGTGCCTA	CCGAGTTGAT	GGCGGGTTG	ACAACTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAACC	CTCTGATTTT	GGCGGAGACC
151	GGCATGGATA	TGGGGGCGGT	ATTCGTCGCT	ACCTGTATCG	CGTCTGCCAT
201	CGGCTGTTTT	GTTATGGGTT	TTGTCGGCAA	CTATCCGATT	GCACTCGCAC
251	CGGGGATGGG	GCTGAATGCC	TATTTACCTT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCCT	GGCAGGTTGC	GTTGGGTGCG	GTGTTTATCT	CCGGTCTGAT
351	TTTCATCCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCATATGGG	TTTGAAAATG	TCGATTGCTG	CCGGTATCGG	TTTGTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCGGC	TTGGGCGGATA	TTCATCAGCC	GTCCGCGTTG	TTGGCACTGT
551	TCGGTTTTGC	CATGGTGGTC	GTATTGGGAC	ATTTCCGCGT	TCAAGGCGCA
601	ATCATCATCA	CCATTTTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAATTT	CACGGCATCA	TCGGCGAAGT	GCCGAGCATT	GCGCCGACTT
701	TTATGCAGAT	GGATTTTAAA	GGGTTGTTTA	CCGTGAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TTTTCCTAGT	CGATCTGTTC	GACAGTACCG	GAACACTGGT
801	CGGTGTATCG	CATCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCCGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CTATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGTGC	CGGGCGTATC
951	GGCAGGCGGG	CGGACAGGTC	TGACGGCGGT	TACCGTCGGC	GTATTGATGC
1001	TCGCCCTGCCT	GATGTTTTCA	CCTTTGGCGA	AAAGTGTTCC	CGCTTTTGCC
1051	ACCGCGCCCG	CCCTGCTTTA	TGTCGGCACG	CAGATGCTCC	GCACTGCGAG
1101	GGACATCGAT	TGGGACGATA	TGACGGAAGC	CGCACCCGCA	TTCCTGACCA
1151	TTGTCTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCTTTTCGGC
1201	TTCATCAGTT	ATGCCGTGGT	TAAACTTTTA	TGCCGCCGCA	CCAAAGACGT
1251	TCCGCCTATG	GTATGGATTG	TTGCCGTATT	GTGGGCACTG	AAATTCTGGT

1301 ATTTGGGCTG A

This corresponds to the amino acid sequence <SEQ ID 344; ORF 097.a>:

a097.pep

```
1  MDTSKQTLTD GIFKLGKANGT TVRTELMAGL TTFLTMCIYIV IVNPLILGET
51  GMDMGAVFVA TCIASAIGCF VMGFVGNYP I ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201 IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFK GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVP PM VWIVAVLWAL KFWYLG*
```

m097/a097 99.3% identity in 436 aa overlap

m097.pep	10	20	30	40	50	60
a097	10	20	30	40	50	60
m097.pep	70	80	90	100	110	120
a097	70	80	90	100	110	120
m097.pep	130	140	150	160	170	180
a097	130	140	150	160	170	180
m097.pep	190	200	210	220	230	240
a097	190	200	210	220	230	240
m097.pep	250	260	270	280	290	300
a097	250	260	270	280	290	300
m097.pep	310	320	330	340	350	360
a097	310	320	330	340	350	360
m097.pep	370	380	390	400	410	420
a097	370	380	390	400	410	420
m097.pep	430					
a097	430					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 345>:

```
g098.seq
  1 ATGACCGCCG ACGGTCTCTT CGTCGCTTTC AACTTCAATA CGTTTGCCGT
 51 TGTGCGAATA TTGATACCAG TACAGCAGGA TGCTGCCGAG GCTGGCGATC
101 AGTTTGTCGG CGATGTCGCG CGCTTCGCTG TCGGGATGGC TTTCGCGTTC
151 GGGATGAACG CAGCCGAGCA TGGACACGCC GGTACGCATC ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAAGTG TCGACTTCT TCAAACTCGC
351 ATTTTGTGC CAAATTAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>:

```
g098.pep
  1 MTADGLFVAF NFNTFAVVRI LIPVQDDAAQ AGDQFVGDDVA RFAVGMAFAF
 51 GMNAAEHGHA GTHVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV CDFFKLAFLC QIRMS*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 347>:

```
m098.seq
  1 ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
 51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
101 AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACTCGC
351 ATTTTGTGC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>:

```
m098.pep
  1 MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDDVA RFTFRMAFTF
 51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV GDDFFKLAFLC QIRMS*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from *N. gonorrhoeae*:

```
m098/g098

      10      20      30      40      50      60
m098.pep MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGYA
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g098      MTADGLFVAFNFNTFAVVRI LIPVQDDAAQAGDQFVGDDVARFAVGMAFAFGMNAAEHGHA
      10      20      30      40      50      60

      70      80      90     100     110     120
m098.pep GTHYVHRMGMCRQAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQVGDFFKLAFLC
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g098      GTHHVHRMGMCRQAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQVCDFFKLAFLC
      70      80      90     100     110     120

m098.pep QIRMSX
|||
g098      QIRMSX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 349>:

```
a098.seq
  1 ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
 51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
```

304

```

101 AGTTTGTCCG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACTCGC
351 ATTTTGTGTC CAAATCAGAA TGTCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 350; ORF 098.a>:

a098.pep

```

1 MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVG DVA RFTFRMAFTF
51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV GDFFKLAFLC QIRMS*

```

m098/a098 100.0% identity in 125 aa overlap

	10	20	30	40	50	60
m098.pep	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVG DVARFTFRMAFTFRMNAAQHGYA					
a098	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVG DVARFTFRMAFTFRMNAAQHGYA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m098.pep	GTHYVHRMGMCRQAFQNFNH TDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
a098	GTHYVHRMGMCRQAFQNFNH TDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
	70	80	90	100	110	120
m098.pep	QIRMSX					
a098	QIRMSX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 351>:

g099.seq

```

1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTGGA
51 GCTGACGGGC AAACGGCAGG CGGGCATTAC TGCCACAGAC ATCGTGTGTTG
101 CACTGACCGA ATTCTTTCGT AAAGAGCGCG TGGTCGGGGC GTTTGTGCGAA
151 TTTTTCGGCG AGGCGCGGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCCATG TTCGCCATCG
251 ACGCGCAAAC TATTGATTAT TTGAACTGA CCGGACGTGA CGACGCGCAG
301 GTGAAATTGG TGGAACCTA CGCCAAACC GCAGGCTTAT GGGCAGGTGG
351 CTTGAAACC GCCGTTTATC CGCGCTTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCCACC
451 GCGGATTTGG CGGCGAAAGG GCTGGCGAAG CCTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCGTGTA
551 CCAATACTTC CAACCCGCGC AACGTTGTCG CCGCCGCACT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGGAATCTA TTTGAAAGAA GCAGGCTTGT
701 TGCCCGAAAT GGAAAACTC GGCTTCGTA TCGTCGCCTT CGCATGTACC
751 ACCTGTAACG GCATGAgcgG CGCGCTcgac CCGAAAATCC AACAAAGAAAT
801 CATCGACCGC GAttgtacg cACCGCCGT ATTGTCAGGC AACCGCAACT
851 TCGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCTTGGTTCG TTGCCTACGC ATTGGCAGGT AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGCGTTG CAGACGGCCG CGAAATCCGC CTGAAAGATA
1001 TCTGGCCGAC AGACGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAACAAT TCCGCGACAT TTATATCCCG ATGTCCGACA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGACCGATG TCCACCTACA
1151 TCCGCCGTCC GCCCTATTGG GAAGGCGCAC TGGCAGGGGA ACGTACATTA
1201 AGAGGTATGC GTCCGCCGCG GATTTTGCCC GACAACATCA CCACCGACCA
1251 CATCTCgcca tCCAATGCGA TTTTGGCCGG cagTGcgcga ggtgaATATT

```

```

1301 TGGCGAAAAT GGGTTTGCCT GAAGAagaCT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
1451 tggcacgcgt tgaaccAGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC GGCGAAGGGC GTGCGGCTGG
1601 CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
1651 AACCTCATCG GCATGGGCGT CTTGCCGCTG CAATTCAAAC CCGGCACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGCTGCGGC CTGACCCTCG TGATTCACCG TAAAAACGGA
1801 GAAACCGTCG AAGTTCCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
1851 ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
1901 TGAAGGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:

```

g099.pep
  1  MLGRASMMRL PDIVGVELTG KRQAGITATD IVLALTEFLR KERVVGAFVE
 51  FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDAQTIDY LKLTGRDDAQ
101  VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNP HARFAT
151  ADLAAKGLAK PYEPPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201  NANRLGLKRK PWKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFAC
251  TCNGMSGALD PKIQQEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFFLASP
301  PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAEYVK
351  PQQFRDIYIP MSDTGTAQKA PSPLYDWRPM STYIRRPYPW EGALAGERTL
401  RGMRPPAILP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
451  RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501  ETYMNRRKQPL I IAGADYQG GSSRDWAAKG VRLAGVEAIA AEGFERIHRT
551  NLIGMGVLPL QFKPGTNRHT LQLDGTETYD VVGERTPRCG LTLVIHRKNG
601  ETVEVPVTCR PDTAEALVY EAGGVLQRFA QDFLEGNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 353>:

```

m099.seq
  1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
 51  GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101  CACTGACCGA GTTCTCGCGC AAAGAACGCG TGGTCGGGGC GTTTGTGCGAA
151  TTCTTCGGCG AGGCGCGGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201  TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
251  ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
301  GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
351  CTTGAAAACC GCCGTTTATC CTCGCGTTTT GAAATTTGAT TTGAGCAGCG
401  TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCAGCC
451  GCCGATTTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
501  CGGCCAAATG CCCGACGGCT CGGTATCAT CATCGCGGATT ACCAGTTGCA
551  CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCACGC
601  AATGCCAACC GTCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651  TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCGGGCCTGT
701  TGCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTCGCCTT CGCCTGCACC
751  ACCTGCAACG GCATGAGTGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
801  CATCGACCGC GATTTGTACG CCACCGCCGT ATTATCAGGC AACCGCAACT
851  TCGACGGCCG TATCCACCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901  CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
951  AAACGACGTA CTCGGCGTTG CAGACGCGCA GGAAATCCGC CTGAAAGACA
1001 TTTGGCCTGC CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTGACGA CCGGCACAGC
1101 GCAAAAAGCA CCCAGTCCGC TGTACGATTG GCGTCCGATG TCCACCTACA
1151 TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
1201 AGAGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251 CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GGCGAGTATT
1301 TGGCGAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTGCGC CAAGGCTCGT
1451 TCGCCCGCGT CGAACCCGAA GGCGAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA

```



```

1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTAGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACCTTATCG GCATGGGCGT GTTGCCGCTG CAGTTCAAAC CCGACACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTGGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTCG TGATTCACCG TAAAAACGGC
1801 GAAACCGTTG AAGTTCCCGT TACCTGCTGC CTCGATACTG CAGAAGAAGT
1851 ATTGATATAT GAAGCCGCGC GCGTGTGCA ACGGTTTGCA CAGGATTTTT
1901 TGAAGGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```

m099.pep
  1 MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
  51 FFGEGARSLS IGDRTISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
 101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
 151 ADLAAKGLAK PYEPPSDGQM PDGSVIIAAI TSCTNTSNPR NVVAAALLAR
 201 NANRLGLKRK PWVKSSFAPG SKVAEILKE AGLLPEMEKL GFGIVAFAC
 251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFSLAP
 301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
 351 PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRPPYVW EGALAGERTL
 401 RGMRLAILLP DNITDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
 451 RGDHLTAQRA TFANPKLFNE MVKNEDGSVR QGSFARVEPE GETMRMWEAI
 501 ETYMNKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHR
 551 NLIGMGVLP LQKPDNRHT LQLDGTETD VVGERTPRCD LTLVIHRKNG
 601 ETVEVPVTCC LDTAEVLVY EAGGVLQRFQ QDFLEGNA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from *N. gonorrhoeae*:

```

m099/g099
      10      20      30      40      50      60
m099.pep MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
          |||||:|||||
g099      MLGRASMMRLPDIVGVELTGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
          10      20      30      40      50      60

      70      80      90      100     110     120
m099.pep IGDRTISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
          |||||:|||||
g099      IGDRTISNMTPEFGATAAMFAIDAQTIDYLKLTGRDDAQVKLVETYAKTAGLWAGGLKT
          70      80      90      100     110     120

      130     140     150     160     170     180
m099.pep AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGSVIIAAI
          |||||:|||||
g099      AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGAIIAAI
          130     140     150     160     170     180

      190     200     210     220     230     240
m099.pep TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAEILKEAGLLPEMEKL
          |||||:|||||
g099      TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAGIYLKEAGLLPEMEKL
          190     200     210     220     230     240

      250     260     270     280     290     300
m099.pep GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFSLAP
          |||||:|||||
g099      GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFSLAP
          250     260     270     280     290     300

      310     320     330     340     350     360

```

307

m099 . pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAVVAEYVKPQQFRDYYVP
g099	PLVVAYALAGSIRFDIENDVLGVADGREIRLKDIWPTDEEIDAIVA EYVKPQQFRDIYIP
	310 320 330 340 350 360
m099 . pep	MFDTGTAQKAPSPLYDWRPMSTYIRRPYWEGALAGERTLRGMRPLAILPDNITTDHLS
g099	MSDTGTAQKAPSPLYDWRPMSTYIRRPYWEGALAGERTLRGMRPPAILPDNITTDHISP
	370 380 390 400 410 420
m099 . pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNE
g099	SNAILAGSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNE
	430 440 450 460 470 480
m099 . pep	QGSFARVEPEGETMRMWEAIEITYMNRKQPLIIAGADYQGSSRDWAAKGVRLAGVEAIV
g099	QGSLARVEPEGQTMRMWEAIEITYMNRKQPLIIAGADYQGSSRDWAAKGVRLAGVEAIA
	490 500 510 520 530 540
m099 . pep	AEGFERIHRTNLIGMGVLPQFKPDTNRHTLQLDGTETD YVVGERTPRCDLTLVIHRKNG
g099	AEGFERIHRTNLIGMGVLPQFKPDTNRHTLQLDGTETD YVVGERTPRCGLTLVIHRKNG
	550 560 570 580 590 600
m099 . pep	ETVEVPVTCCLDTAEVLVYEAGGVLRFAQDFLEGNAAX
g099	ETVEVPVTCRDPDTAEALVYEAGGVLRFAQDFLEGNAAX
	610 620 630 640

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 355>:

a099 . seq

1	ATGCTGGGAC	GCGCGTCCAT	GATGCGCCTG	CCCGATATTG	TCGGCGTTGA
51	GCTGAACGGC	AAACGGAAGG	CGGGCATTAC	GGCGACGGAT	ATTGTGTGTTG
101	CACGTACCGA	GTTTCTGCGC	AAAGAACGCG	TGGTCGGGGC	GTTTGTGCGAA
151	TTCTTCGGCG	AGGGCGCGAG	AAGCCTGTCT	ATCGGCGACC	GCGCGACCAT
201	TTCCAACATG	ACGCCGAGT	TCGGCGCGAC	TGCCGCGATG	TTCGCTATTG
251	ATGAGCAAAC	CATTGATTAT	TTGAAACTGA	CCGGACGCGA	CGACGCGCAG
301	GTGAAATTGG	TGGAACCTA	CGCCAAACC	GCAGGCTTGT	GGGCAGATGC
351	CTTGAAAACC	GCCGTTTATC	CGCGCGTTT	GAAATTTGAT	TTGAGCAGCG
401	TAACGCGCAA	TATGGCAGGC	CCGAGCAACC	CGCACGCGCG	TTTTGCGACC
451	GCCGATTTGG	CCGGCAAAGG	CTTGGCTAAA	CCTTACGAAG	AGCCTTCAGA
501	CGGCCAAATG	CCTGACGGTG	CAGTGATTAT	TGCCGCGATT	ACTTCCTGTA
551	CCAATACTTC	CAATCCGCGC	AACGTTGTG	CCGCCGCGCT	GTTGGCACGC
601	AATGCCAACC	GCCTCGGCTT	GCAACGCAAA	CCTTGGGTGA	AATCTTCGTT
651	TGCCCCGGGT	TCAAAAGTAG	CCGAAATCTA	TTTGAAAGAA	GCAGATCTGC
701	TGCCCGAAAT	GGAAAACTC	GGCTTCGGTA	TCGTTGCCTT	CGCATGTACC
751	ACCTGTAACG	GCATGAGCGG	CGCGCTGGAT	CCGAAAATCC	AGAAAGAAAT
801	CATCGACCGC	GATTTGTACG	CCACCGCCGT	ATTGTCAGGC	AACCGCAACT
851	TTGACGGCCG	TATCCATCCG	TATGCGAAAC	AGGCTTTCCT	CGCTTCGCCT
901	CCGTTGGTCG	TTGCCTACGC	GCTGGCAGGC	AGCATCCGTT	TCGATATTGA
951	AAACGACGTA	CTCGGCGTTG	CAGACGGCAA	AGAAATCCGC	CTGAAAGACA
1001	TTTGGCCTAC	CGATGAAGAA	ATCGATGCCA	TCGTTGCCGA	ATATGTGAAA
1051	CCGACGCAAT	TTCGCGACGT	TTATATCCCG	ATGTTGACA	CCGGCACAGC
1101	GCAAAAAGCA	CCAAGCCCGC	TGTACGACTG	GCGTCCAATG	TCTACCTATA
1151	TCCGCCGCC	ACCTTACTGG	GAAGGCGCAC	TGGCAGGGGA	ACGCACATTA
1201	AGCGGTATGC	GTCCGCTGGC	GATTTTGCCC	GACAACATCA	CCACCGACCA

```

1251 TCTCTCGCCA TCCAATGCGA TTTTGGCAAG CAGTGCCGCA GGCGAATATT
1301 TGGCAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTTCGC
1451 TGGCACGCGT TGAACCCGAA GGCCAAACCA TCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGCGCGGA
1551 CTACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCTCG
1601 CCGGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACTTGATCG GTATGGGCGT GTTGCCGCTG CAGTTCAAAC CGGGTACCAA
1701 CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTTG TGATTCACCG TAAAAACGGC
1801 GAGACCGTCG AAGTCCCAT TACCTGCCGC CTCGATACCG CAGAAGAAGT
1851 GTTGGTATAT GAAGCCGGTG GCGTATTGCA ACGGTTTGCA CAGGATTTTT
1901 TGAAGGGAA CGCGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 356; ORF 099.a>:

a099.pep

```

1 MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRTISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAGKGLAK PYEEPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLQRK PWVKSSFAPG SKVAEIYLKE ADLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPTDEE IDAIVAEYVK
351 PQFRDVYIP MEDTGTAQKA PSPLYDWRPM STYIRPPYVW EGALAGERTL
401 SGMRLAILP DNITTDHLSN SNAILASSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMRNRKQPL IIIAGADYQ GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
551 NLIGMGVLP QFKPGTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
601 ETVEVPITCR LDTAEEVLVY EAGGVLRFA QDFLEGNA*

```

m099/a099 97.5% identity in 639 aa overlap

	10	20	30	40	50	60
m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
a099	MLGRASMMRLPDIVGVELNGKRKAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m099.pep	IGDRATISNMTPEFGATAAMFAIDEQTIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
a099	IGDRATISNMTPEFGATAAMFAIDEQTIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
	70	80	90	100	110	120
	130	140	150	160	170	180
m099.pep	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI					
a099	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAGKGLAKPYEEPSDGQMPDGAIIAAI					
	130	140	150	160	170	180
	190	200	210	220	230	240
m099.pep	TSCTNTSNPRNVVAAALLARNANRLGLKRKPWKSSFAPGSKVAEIYLKEAGLLPEMEKL					
a099	TSCTNTSNPRNVVAAALLARNANRLGLQRKPWKSSFAPGSKVAEIYLKEADLLPEMEKL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m099.pep	GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFASP					
a099	GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFASP					
	250	260	270	280	290	300
	310	320	330	340	350	360
m099.pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIPADEEIDAIVAEYVKPQFRDVYVP					

a099	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPTDEEIDAIVA EYVKPQQFRDVYIP	310	320	330	340	350	360
m099.pep	MFDTGTAQKAPSPLYDWRPMS TYIRRPYEWEGALAGERTLRGMRPLAILPDNITTDHLS P	370	380	390	400	410	420
a099	MFDTGTAQKAPSPLYDWRPMS TYIRRPYEWEGALAGERTLSGMRPLAILPDNITTDHLS P	370	380	390	400	410	420
m099.pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR	430	440	450	460	470	480
a099	SNAILASSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR	430	440	450	460	470	480
m099.pep	QGSFARVEPEGETMRMWEA IETYMNRKQPLII IAGADYQGSSRDWA AKGVRLAGVEAIV	490	500	510	520	530	540
a099	QGSLARVEPEGQTRMWEA IETYMNRKQPLII IAGADYQGSSRDWA AKGVRLAGVEAIV	490	500	510	520	530	540
m099.pep	AEGFERIHRTNLIGMGVLP LQFKPDTNRHTLQLDGTET YDVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
a099	AEGFERIHRTNLIGMGVLP LQFKPGTNRHTLQLDGTET YDVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
m099.pep	ETVEVPVTCCLDTAE EVLVYEAGGV LQRFAQDFLEGNAAX	610	620	630	640		
a099	ETVEVPITCRLDTA E EVLVYEAGGV LQRFAQDFLEGNAAX	610	620	630	640		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 357>:

g102.seq

1	AtgtCCGCCA	AAactccgtc	gctcttcggc	ggcgcatga	Ttatcgccgg
51	gaaggttaat	ggcgcaGgta	tgttcccaa	ccccaccgcc	aacttggggg
101	acgggttaac	aggtctcgct	attgtgctg	tgtacacctg	gtttccattc
151	tccgtcgcg	ccctcatgat	tttggaaagc	aacacccata	acCCCcgagg
201	ggcaAGtttt	gacaccATGg	tcAAagacct	gtcgcGaCGc	ggctggaaca
251	tcatcaacgg	catcgccgct	gctttggTCc	gatacggctc	gacctacgcg
301	tacattttag	tcgcgggtga	cttGACCGCC	AAAGGCAtcg	GCAGCGAGT
351	AGGCGGCAAA	ATTTCgctca	CCGTCGGACA	actcgtcttc	tTCGGCATCC
401	TCGCCTTTTG	CGTATGGGCA	TCCGCACGCT	TGGTCGACCG	CTTTACCGGC
451	GTCTCATCG	GCGGCATGTT	ATTAACCTTT	ATTTGGGCCAA	CGGGCGGCTT
501	GGTTGCCAGT	GCCAAACCGT	CCGTCCTCTT	CGACACCCAA	GCCCCGCTCG
551	GCACCGGGAT	TCGGATTACG	GCCGCCACCG	CCCTGCCCGT	CTGCCCTAAGT
601	TCCTTCGGCT	TCCACGGCAA	CGTTTCCAGC	CTGCTCAAAT	ACTTTTAAGG
651	CGACGcgcc	aaagtGgCGA	ATTCaatctg	gGcaggtaca	ttggTTGCCT
701	tggtaattta	cgctcTCTgg	caaaccgcc	tCcaaagcaa	ccTGCcgcg
751	aacgagctcg	cCCCcgtagt	tgccgcggag	aggcaactCT	CCGTCTTGaa
801	tgaaacctTG	tccaaattcg	cccaaaccgg	cgatatggat	aAaatattgt
851	ccctatttcc	tcatatggca	atcgcaacct	cttttttagg	cgtTAACctta
901	ggctgtgttg	acaacatcgc	cgacatcttc	aaatggaacg	acagttatgtc
951	cggcgggggc	accaaaaccg	tcgcgctgaa	cttcttgccg	CCCCtgattt
1001	cgtgctgct	ctcccccaac	ggctttctta	ccgcatattgg	tcgctccggc
1051	ctggcgggcaa	ccgtctggga	ccaagGcacc	atccccgcc	tgtctgctca
1101	cgtttcccc	caaaaaattG	gcGcaggcaa	gactataAa	gtttaCGGCG
1151	gcttgtggct	Gatcttagtc	ttcctttttc	gcctgcgcaa	catcgcgcga
1201	CAGGTATTGA	GcaAatTGa	ActcgtCccc	GTATTTAAG	GATTA

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>:

q102.ppt

```

1  MSAKTPSLFG  GAMIIAGKVI  GAGMFPNPTA  NLGDGLIGSL  IVLLYTWFPP
51  SSGALMILEV  NTHNPRGASF  DTMVKDLLGR  GWNIINGIAV  ALVLYGSTYA
101 YILVGGDLTA  KGIGSAVGGK  ISLTVGQLVF  FGILAFVCWA  SARLVDREFTG
151 VLIGGMVLTF  IWATGGLVAD  AKPSVLFDTQ  APVGTGYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWAGT  LVALVIYVLW  QTAIQSNLPR
251 NEFAPVIAAE  RQLSVLNETL  SKFAQTGDM  KILSLFPYMA  IATSFLGVTL
301 GLFDNIADIF  KWNDMSMGRG  TKTVALNFLP  PLISWLLLPT  GFFTAIGASG
351 LAATVWDQGI  IPAMLLYVSP  QKIGAGKTYK  VYGGLWMLLV  FLFGIANIAA
401 QVLSQMELV  P  VFKG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 359>:

```

m102.seq
1  ATGCCCAACA  AAACCCCTTC  ACTGTTTCGGC  GGC CGCATGA  TTATCGCCGG
51  CACGGTCATC  GGC CGCAGGCA  TGCTCGCCAA  CCCGACCGCC  ACATCCGGCG
101 TATGGTTTAC  CGGCTCGCTG  GCCGTGTTGC  TGTACACCTG  GTTTTCTATG
151 CTTTCCAGCG  GCCTGATGAT  TTTGGAAGTC  AACACCCATT  ATCCGCACGG
201 CGCAAGTTTC  GACACGATGG  TCAAAGACCT  GCTCGGACGC  GGCTGGAACA
251 TCATCAACGG  CATCGCCGTC  GCCTTCGTTT  TATACCTGCT  TACTTACGCT
301 TATATCTTCG  TCGGCGGCGA  CCTGACCGCC  AAAGGCTTAG  GCAGCGCGGC
351 AGGCGGCGAC  GTTCACTCA  CCGTCGGACA  ACTCGTCTTC  TTCGGCATCC
401 TCGCCTTTTG  CGTATGGGCA  TCCGCACGCT  TGGTCGACCG  CTTACCGGC
451 GTCCTTATCG  GCGGCATGGT  ATTGACCTTT  ATTTGGGCGG  CCGGCGGGCT
501 GATTGCCGAT  GCCAAGCCGT  CCGTCTCTTT  CGATACCCAA  GCCCCGCGCG
551 GCACAACTA  CTGGATTAC  GCCGCCACCG  CCCTGCCCGT  CTGCCTCGCT
601 TCCTTCGGCT  TCCACGGCAA  CGTCTCCAGC  CTGCTCAAAT  ACTTTAAAGG
651 CGACGCGCCC  AAAGTGGCTA  AATCCATCTG  GACGGGCACA  CTGATTGCGC
701 TGGTAATTTA  CGTCTCTGG  CAAACCGCCA  TCCAAGGCAA  CCTGCCGCGC
751 AACGAGTTTC  CCCCCTCAT  CGCCGCCGAA  GGGCAAGTCT  CGTCTCAT
801 CGAAACCTG  TCCAAATTCG  CCCAAACCGG  CAATATGGAC  AAAATATTGT
851 CCCTGTTTTC  CTATATGGCG  ATCGCCACCT  CGTTTTTAGG  CGTAACGCTC
901 GGAATCTTCG  ACTACATCGC  CGACATCTTC  AAATGGAACG  ACAGCATCTC
951 CGGCCGCACC  AAAACCGCG  CGCTGACCTT  CCTGCCGCCC  CTGATTTCTT
1001 GCCTGCTCTT  CCCACCGGC  TTCGTTACCG  CCATCGGCTA  CGTCGGCCTG
1051 GCGGCAACCG  TCTGGACAGG  CATCATCCCC  GCCATGCTGC  TCTACCGTTC
1101 GCGCAAAAAA  TTCGGCGCAG  GCAAAACCTA  TAAAGTTTAC  GCGGCGTTGT
1151 GGCTGATGGT  TTGGGTCTTC  CTTTTCGGCA  TCGTCAACAT  CGCCGCACAG
1201 GTATTGAGCC  AAATGGAAC  CGTCCCCGTA  TTTAAAGGAT  AA

```

This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>:

```

m102.pep..
1  MPNKTPSLFG  GAMIIAGTVI  GAGMLANPTA  TSGVWFTGSL  AVLLYTWFSM
51  LSSGLMILEV  NTHYPHGASF  DTMVKDLLGR  GWNIINGIAV  AFVLYLLTYA
101 YIFVGGDLTA  KGLGSAAGGD  VSLTVGQLVF  FGILAFVCWA  SARLVDREFTG
151 VLIGGMVLTF  IWAAGGLIAD  AKPSVLFDTQ  APAGTNYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWTGT  LIALVIYVLW  QTAIQGNLPR
251 NEFAPVIAAE  QQSVLIETL  SKFAQTGND  KILSLFSYMA  IATSFLGVTL
301 GLFDYIADIF  KWNDISISRT  KTAALTFLPP  LISCLLFPPT  FVTAIGYVGL
351 AATVWTGIIP  AMLYRSRKK  FGAGKTYKVY  GGLWLMVWVF  LFGIVNIAAQ
401 VLSQMELV  P  FKG*

```

m102/g102 86.0% identity in 415 aa overlap

	10	20	30	40	50	60
m102.pep	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
g102	MSAKTPSLFGGAMIIAGKVIAGMFPNPTANLGDGLIGSLIVLLYTWFPFSSGALMILEV					
	10	20	30	40	50	60
m102.pep	NTHYPHGASFDTMVKDLLGRGWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD					
g102	NTHNPRGASFDTMVKDLLGRGWNIINGIAVALVLYGSTYAYILVGGDLTAKGIGSAVGGK					
	70	80	90	100	110	120
m102.pep	VSLTVGQLVFFGILAFVCWASARLVDREFTGVLLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
g102	ISLTVGQLVFFGILAFVCWASARLVDREFTGVLLIGGMVLTFIWAAGGLVADAKPSVLFDTQ					
	130	140	150	160	170	180
m102.pep	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTIALLVIYVLW					
g102	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTIALLVIYVLW					
	190	200	210	220	230	240

a102.seq

1	ATGCCACCA	AAACCCCTTC	ACTGTTCCGC	GGCGCGATGA	TTATCGCCGG
51	CACGNTCATC	GGCGCAGGTA	TGCTCGCCAA	CCCGACCGCC	ACATCCGGCG
101	TATGTTTAC	CGGCTCGCTG	GCCGTTGTC	TGTACACCTG	GTTTTCCATG
151	CTCTCCAGCG	GCGTGTGATG	TTTGGAAAGT	AACACCCACT	ACCCCGACGG
201	CGCGANCTTC	GACACCATGG	TTAAAGACCT	GCTCGGACGG	AGCTGGAACG
251	TCATCAACGG	CATCGCCGTC	GCCTTCGTTT	TATACCTGCT	TACTTACGCT
301	TATATCTTCG	TCGGCGCGCA	CCTGACCGCC	AAAGGCTTAG	GCAGCGCGCG
351	AGGCGGCAAT	GTTTCACTCA	CCGTCCGACA	ACTCGTCTTC	TTCCGGCAATC
401	TCGCCTTTTG	CGTATGGGCA	TCCGCACGCT	TGGTCGACCG	ATTACACGAG
451	GTCCTCATCG	GCGGCATGGT	ATTAACCTTT	ATTTGGGCAA	CCGGCGGCCT
501	GATTGCGCAT	GCCAAACTGC	CCGTCCCTCT	CGACACCCAA	GCCCCTACCG
551	GCACCAACTA	TCGGATTTAT	GTCGCCACCG	CCCTGCCCGT	TCGCCTTGCG
601	TCATTTCGGT	CCTGACGGCA	CGTCTCCAGC	CTGCTCAAAT	ACTTTAAAGG
651	CGACGCGCCC	AAAGTGGCTA	AATCCATCTG	GACGGGCACA	CTGATTGCGC
701	TGGTAATTTA	CGTCCTCTGG	CAAACCGCCA	TCCAANGCAA	CCTGCCGCGC
751	AACGAGTTCG	CCCCCGTGAT	TGCCGCCGAA	GGGCAAGTCT	CCGTCNTGAT
801	TGAAACCCCTG	TCCAATTCG	CCCAACCCGG	CAATATGGAC	AAATATTTGT
851	CCCTGTTTTT	CTATATGGCG	ATCGCCACCT	CGTTTTTtag	CGTAACGCTC
901	GGACTCTTCG	ACTACATCGC	CGACATCTTC	AAATGGAACG	ACAGCGTGTC
951	CGGCCGCAAC	AAAACCGCCG	CGCTGACCTT	CCTGCCGCCT	NTAATTTTCT
1001	GCTTGCTCTT	CCCCACCGCG	TTTGTTACCG	CCATCGGNTA	CGTCCGCTGT
1051	GCGGCAACCCG	TCTGGACAGG	CATCATCCCC	GCGATGCTGC	TNTACCGTTT
1101	GCGCAAAAAA	TTCCGGCGCAG	GCAAAACCTA	TAAAGTTTAC	GGCGGCTTGT
1151	GGCTGATGGT	TTGGGTCTTC	CTTTTCGGCA	TCNTCAACAT	CGCCGCACAN
1201	GTATTGAGCC	AAATGGAAC	CGTCCCCGTA	TTTAAAGGAT	AA
1202					

This corresponds to the amino acid sequence <SEQ ID 362; ORF 102.a>:

a102.pep

1	MPTKTPSLFG	GAMIIAGTXI	GAGMLANPTA	TSGVWFTGSL	AVLLYTWFMS
51	LSSGLMILEV	NTHYPHGAXF	DTMVKDLLGR	SWNIINGIAV	AFVLYLLTYA
101	YIFVGGDLTA	KGLGSAAGGN	VSLTVGQDLV	FGILAFCVWA	SARLVDRFST
151	VLIGGMVLTF	IWATGGLIAD	AKLPVLFDTQ	APTGTNIWIY	VATALPVCLA
201	SFGFHGNVSS	LLKYFKGDAP	KVAKSIWTGT	LIALVIYVLW	QTAIQXNLPR
251	NEFAPVIAAE	GQVSXVIETL	SKFAQTGNMD	KILSLFSYMA	IATSFGLVTL
301	GLFDYIADIF	KWNDSVSGRT	KTAALTFLPP	XISCLLPFTG	FVTAIGYVGL
351	AATVWGTGII	AMLLYRSRKK	FGAGKTYKVY	GGLWLMVWVF	LFGIXNIAAX
401	ALTSOMELVP	EKG*			

m102 / a102 95.9% identity in 413 aa overlap

312

	10	20	30	40	50	60
m102.pep	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
a102	MPTKTPSLFGGAMIIAGTXIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m102.pep	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD					
a102	NTHYPHGAXFDTMVKDLLGRSWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGN					
	70	80	90	100	110	120
	130	140	150	160	170	180
m102.pep	VSLTVGQLVFFGILAFCVWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
a102	VSLTVGQLVFFGILAFCVWASARLVDRFTSVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
	130	140	150	160	170	180
	190	200	210	220	230	240
m102.pep	APAGTNYWIYAATALPVCCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
a102	APTGTNYWIYVATALPVCCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
	190	200	210	220	230	240
	250	260	270	280	290	300
m102.pep	QTAIQGNLPRNEFAPVIAAEGQVSVLIELTSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
a102	QTAIQXNLPRNEFAPVIAAEGQVSXIELTSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
	250	260	270	280	290	300
	310	320	330	340	350	360
m102.pep	GLFDYIADIFKWNDISGRKTAAALFLPPLISCLLFPTGFVTAIGYVGLAATVWTGIIP					
a102	GLFDYIADIFKWNDISVGRKTAAALFLPPXISCLLFPTGFVTAIGYVGLAATVWTGIIP					
	310	320	330	340	350	360
	370	380	390	400	410	
m102.pep	AMLLYRSRKKFGAGKTYKVYGGWLWMVWVFLFGIVNIAAQVLSQMELVPVFKGX					
a102	AMLLYRSRKKFGAGKTYKVYGGWLWMVWVFLFGIXNIAAXVLSQMELVPVFKGX					
	370	380	390	400	410	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 363>:

```

g105.seq
1  Atgtccgcag aaaCATACAC acAAAtcggc tGGgtaggct taggGcaaat
51  gGgtctgcct atgGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101 TATACAACCG CTCGCCCCGAC AAAACTGCCC CCATCTCegc CAAAGGAGCA
151 AAAGTTTACG GCagcACCGC CGAACTCGTC CGCGCCTGCC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351 TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
401 TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451 ATATTTTCCC TTGTGGGCAA AAAAACCTTC CATTTGCGCG ATGTGGGCAA
501 AGGCTCGGGC GCGAAACTCG TCTTGAATC GCTCTTAGGC ATTTTCGGCG
551 AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601 GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCTATGTT
651 TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CCTGCCTTTG
701 CACTCAAACA CGCTTCCAAA GAcctTAACC TCGccgtcAA AGAGCTTGAA
751 CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801 CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851 TGAAATTGGC AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 364; ORF 105.ng>:

```
g105.pep
  1  MSAETYTQIG WVGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGSTAELV RACPVIIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTG HFGDVGKGS AKLVLSLLG IFGEAYSEAM LMARQFGIDT
201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 365>:

```
m105.seq
  1  ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGaTAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAm ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAaGC TCGCCGTCa AGCACTTGTC GAAGCGCAGm GaCAGTTTGC
351  CGAAGCACCC GTTTCCGGAT CGGTCGGGCC CGCCACCAAC GGCACGCTGC
401  TGATTCTGTT CGGCGGCAGC GAaCCGtTTT AAACCCGCTG CAAAAAATAT
451  TTTCCCTCGT CGGCAAAAA ACCTTCCATT TCGGCGATGT CGGCAAAGGT
501  TCGGGCGCGA AACTCGTCTT GAACCTCGTC TTGGGCATTT TCGGCGAaCG
551  TAcAGCGAAs GmTgCTGATG GCGCGGCAGT TCGGCATCGA TACCGACACC
601  ATCGTCGAAG CCATCGGsGA CTCGGCAATG GACTCGCCCA TGTTCCAAAC
651  CAAAAAATCC CTGTGGGCAA ACCGCGAATT CCCGmCCGmC TTCGCCCTCA
701  AACACGCCTC CAAAGACCTC AACCTCGCCG TCAAAGAGCT TGAACAGGCA
751  GGCAACACCC TGCCCGCCGT CGAAACCGTT GCTGCCAGCT ACCGCAAAGC
801  AGTCGAAGCC GGCTACGGGA CACAGGACGT TTCCGGCGTT TACCTGAAAC
851  TGGCAGAACA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:

```
m105.pep
  1  MSANEYAQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51  KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGX IIVNMSTISP
101  TEKLAVKALV EAQRQFAEAP VSGSVGPATN GTLLILFGGS EPFXTRCKKY
151  FPSSAKKPSI SAMSAKVRAR NSSXTRSWAF SANVQRXXLM ARQFGIDTDT
201  IVEAIGDSAM DSPMFQTKKS LWANREFPXX FALKHASKDL NLAVKELEQA
251  GNTLPAVETV AASYRKAVEA GYGTQDVSGV YLKLAEH
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 105 shows 79.9% identity over a 289 aa overlap with a predicted ORF (ORF 105.ng) from *N. gonorrhoeae*:

m105/g105

	10	20	30	40	50	60
g105.pep	MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV					
m105	MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
	10	20	30	40	50	60
g105.pep	RACPVIIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA					
m105	RDYPVIFLMVSDYAAVCDILNGVRDGLAGXIIIVNMSTISPTTEKLAVKALVEAQR-QFAEA					
	70	80	90	100	110	120
g105.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSAGKLVLSLLG					
m105	PVSGSVGPATNGTLLILFGGSEPFXTCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW					
	120	130	140	150	160	170

314

	190	200	210	220	230	240
g105.pep	IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK					
	: :					
m105	AFSANVQRXXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXXFALKHASK					
	180	190	200	210	220	230

	250	260	270	280	289
g105.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEH				
m105	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGTQDVSGVYLKLAEH				
	240	250	260	270	280

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 367>:

a105.seq

```

1  ATGTCCGCAA  ACGAATACAC  ACAAATCGGC  TGGATAGGCT  TAGGGCAAAT
51  GGGCTGCGCT  ATGGTAACGC  GGCTCTTGA  CGGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCGAC  AAAACTGCCC  CCATCTCCGC  CAAAGGCGCA
151 AAAGTTTACG  GCAACACCGC  CGAACTCGTC  CGCGACTATC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CCGCCGTGTG  CGACATCCTG  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCCG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTCGG  GCCCGCCACC  AACGGCACGC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TCGTCGGCAA  AAAACCTTC  CATTTCGGCG  ATGTCGGCAA
501 AGGTCGGGC  GCGAACTCG  TCTTGAATC  GCTCTTGGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAGCGATG  CTGATGGCGC  GGCAGTTCCG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGCGGCTCG  GCAATGGACT  CGCCCATGTT
651 CCAAACCAAA  AAATCCCTGT  GGGCAAACCG  CGAATTCCCA  CCCGCCTTCG
701 CCCTCAAAAC  CGCCTCCAAA  GACCTCAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCTGCC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGCTTTACC
851 TGAAATTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 368; ORF 105.a>:

a105.pep

```

1  MSANEYTQIG  WIGLGQMLP  MVTRLLDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGNTAELV  RDYPVIFLMV  SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP
101 TENLAVKALV  EAAGGQFAEA  PVSGSVGPAT  NGTLLILFGG  SEAVLNPLQK
151 IFSLVGKKT  FFGDVGKGS  AKLVLSLLG  IFGEAYSEAM  LMARQFGIDT
201 DTIVEAIGGS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251 QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*

```

m105/a105 96.5% identity in 289 aa overlap

	10	20	30	40	50	60
m105.pep	MSANEYAQIGWIGLGQMLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
	:					
a105	MSANEYTQIGWIGLGQMLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
	10	20	30	40	50	60

	70	80	90	100	110	119
m105.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISP TENLAVKALVEAAG-QFAEA					
a105	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISP TENLAVKALVEAAGGQFAEA					
	70	80	90	100	110	120

	120	130	140	150	160	170	179
m105.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSAGKLVLSLLG						
a105	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSAGKLVLSLLG						
	130	140	150	160	170	180	

	180	190	200	210	220	230
--	-----	-----	-----	-----	-----	-----

315

```

m105.pep  IFGDV-QRXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFFXAFALKHASK
a105      IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
           190      200      210      220      230      240

m105.pep  240      250      260      270      280
DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
a105      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
           250      260      270      280      290

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 369>:

g105-1.seq

```

1  ATGTCCGCAG  AAACATACAC  ACAAATCGGC  TGGGTAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGGTAACGC  GGCTCTTGGA  CGGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCGAC  AAAACTGCCC  CCATCTCCGC  CAAAGGAGCA
151 AAAGTTTACG  GCAGCACCGC  CGAACTCGTC  CGCGCCTGCC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CCGCCGTGTG  CGACATCTCG  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCGG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTCGG  ACCCGCCACC  AACGGCACAC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TTGTCGGCAA  AAAAACCTTC  CATTTTCGGC  ATGTCGGCAA
501 AGGCTCGGGC  GCGAAACTCG  TCTTGAATC  GCTCTTAGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAGCGATG  CTGATGGCGC  GGCAGTTCCG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGCGGCTCG  GCAATGGACT  CGCCTATGTT
651 TCAAACAAAA  AAATCACTAT  GGGCAAACCG  TGAGTTCCCC  CCTGCCTTTG
701 CACTCAAACA  CGCTTCCAAA  GACCTTAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCCTGCC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTT  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGGTTTACC
851 TGAATTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>:

g105-1.pep

```

1  MSAETYQIG  WVGLGQMLP  MVTRLDDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGYSTAELV  RACPVIIFLMV  SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP
101 TENLAVKALV  EAAGGQFAEA  PVSQSVGPAT  NGTLLILFGG  SEAVLNPLQK
151 IFSLVGKTF  HFGDVGKSG  AKLVNLSLG  IFGEAYSEAM  LMARQFGIDT
201 DTIVEAIGGS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251 QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 371>:

m105-1.seq

```

1  ATGTCCGCAA  ACGAATACGC  ACAAATCGGC  TGGATAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGGTAACGC  GGCTCTTGGA  CGGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCGAC  AAAACTGCCC  CCATCTCCGC  CAAAGGCGCA
151 AAAGTTTACG  GCAACACCGC  CGAACTCGTC  CGCGACTATC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CCGCCGTGTG  CGACATCTCG  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCGG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTCGG  GCCCGCCACC  AACGGCACGC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TCGTCGGCAA  AAAAACCTTC  CATTTTCGGC  ATGTCGGCAA
501 AGGTTCGGGC  GCGAAACTCG  TCTTGAATC  GCTCTTGGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAnCGATG  CTGATGGCGC  GGCAGTTCCG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGSGACTCG  GCAATGGACT  CGCCCATGTT
651 CCAAACAAA  AAATCCCTGT  GGGCAAACCG  CGAATTCCCG  CCGCCTTCG
701 CACTCAAACA  CGCTTCCAAA  GACCTCAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCCTGCC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGGTTTACC
851 TGAAACTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 372; ORF 105-1>:

m105-1.pep

```

1  MSANEYAQIG  WIGLGQMLP  MVTRLDDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGYNSTAELV  RDYPVIFLMV  SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP

```

316

```

101  TENLAVKALV  EAAGGQFAEA  PVSGSVGPAT  NGTLLILFGG  SEAVLNPLQK
151  IFSLVGKKTf  HFGDVGKGSg  AKLVLSLLG  IFGEAYSEXM  LMARQFGIDT
201  DTIVEAIGDS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251  QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*

```

m105-1/g105-1 96.9% identity in 289 aa overlap

```

          10      20      30      40      50      60
m105-1.pep  MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVG VYNRSPDKTAPISAKGAKVYGNTAELV
          |||: ||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g105-1      MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVG VYNRSPDKTAPISAKGAKVYGSTAELV
          10      20      30      40      50      60

          70      80      90     100     110     120
m105-1.pep  RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
          |  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g105-1      RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
          70      80      90     100     110     120

          130     140     150     160     170     180
m105-1.pep  PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTfHFGDVGKGSgAKLVLSLLG
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g105-1      PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTfHFGDVGKGSgAKLVLSLLG
          130     140     150     160     170     180

          190     200     210     220     230     240
m105-1.pep  IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g105-1      IFGEAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
          190     200     210     220     230     240

          250     260     270     280     290
m105-1.pep  DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSgVYLKLAEHX
          |||:|||||:|||||:|||||:|||||:|||||:|||||
g105-1      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSgVYLKLAEHX
          250     260     270     280     290

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 373>:

a105-1.seq

```

1   ATGTCCGCAA  ACGAATACAC  ACAAATCGGC  TGGATAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGGTAACGC  GGCTCTTGGA  CGGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCGAC  AAAACTGCCC  CCATCTCCGC  CAAAGGCGCA
151 AAAGTTTACG  GCAACACCGC  CGAACTCGTC  CGCGACTATC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CCGCCGTGTG  CGACATCTCG  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCGG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCGGTTTCCG  GATCGGTCGG  GCCCGCCACC  AACGGCACGC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TCGTCGGCAA  AAAAACCTTC  CATTTCCGGC  ATGTCGGCAA
501 AGGTTCGGGC  GCGAAACTCG  TCTTGAACCT  GCTCTTGGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAGCGATG  CTGATGGCGC  GGCAGTTCGG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGCGGCTCG  GCAATGGACT  CGCCCATGTT
651 CCAAACCAAA  AAATCCCTGT  GGGCAAACCG  CGAATCCCA  CCCGCCTTCG
701 CCCTCAACA  CGCTCCAAA  GACCTCAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCCTGCC  CGCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GGCGTTTACC
851 TGAAATTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:

a105-1.pep

```

1   MSANEYTQIG  WIGLGQMGLP  MVTRLLDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGNTAELV  RDYPVIFLMV  SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP
101 TENLAVKALV  EAAGGQFAEA  PVSGSVGPAT  NGTLLILFGG  SEAVLNPLQK
151 IFSLVGKKTf  HFGDVGKGSg  AKLVLSLLG  IFGEAYSEAM  LMARQFGIDT
201 DTIVEAIGGS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251 QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*

```

a105-1/m105-1 99.0% identity in 289 aa overlap

```

          10      20      30      40      50      60
a105-1.pep  MSANEYTQIGWIGLGQMGLPMVTRLLDGGIEVG VYNRSPDKTAPISAKGAKVYGNTAELV
          |||:||||:|||||:|||||:|||||:|||||:|||||:|||||

```

m105-1	MSANEYAQIGWIGLGMGLPMVTRLRLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
	10 20 30 40 50 60
a105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
m105-1	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
a105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKTFHFQDVGKSGAKLVLSLLG
	130 140 150 160 170 180
m105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKTFHFQDVGKSGAKLVLSLLG
	130 140 150 160 170 180
a105-1.pep	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
m105-1	IFGEAYSEXMLMARQFGIDTDITVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
a105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
	250 260 270 280 290
m105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
	250 260 270 280 290

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 375>:

```

g107.seq
1  ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGGTTG CCGATGCCAA
51  ACCGTCCGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
101 TTTACGCCGC CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
201 GgCGAaATCc atctggGcag gtacattggT TGCCTtggtta atttacgtcc
251 TCTggcaaac cgccatCcaa agcaaccTGC cgcgcaacga gttcgcCCCC
301 gtgattgccg ccgagaggca actCTCCGTC CTgaatgaaa cccTGtccaa
351 attcgcccaa accggcgata tggataAaat attgtccccta ttccctaca
401 tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
451 atcgccggac atcttcaaat ggaacgacag tatgtccggg cggcaccaaa
501 accgtcgcgc tga

```

This corresponds to the amino acid sequence <SEQ ID 376; ORF 107.ng>:

```

g107.pep
1  MVLTFIWATG GLVADAKPSV LFDTQAPVGT GYWIYAATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP
101 VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
151 IAGHLQMERQ YVRAAPKPSR *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 377>:

```

m107.seq
1  ATGGTATTGA CCTTTATTTG GCGGCGCGGC GGGCTGATTG CCGATGCCAA
51  GCCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
101 TTTACGCCGs CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGCTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCCTTGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
351 ATTCGCCCCA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGgCCG CACCAAAACC
501 GCCGCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>:

```

m107.pep.
1  MVLTFIWAAG GLIADAKPSV LFDTQAPAGT NYWIYAXTAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGLIALV IYVLWQTAIQ GNLPRNEFAP

```

101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151 IAHLMERQH LRAAPKPPR*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng) from *N. gonorrhoeae*:

m107/g107

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	: : : : :					
g107	MVLTFIWATGGLVADAKPSVLFDTQAPVGTGYWIYAATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	: : : : :					
g107	KGDAPKVAKSIWAGTLVALVIYVLWQTAIQSNLPRNEFAPVIAAERQLSVLNETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSF LGVTLGLFDYIA-HLMERQHLRAAPKPPR					
	: : : : :					
g107	TGDMDKILSLFPYMAIATSF LGVTLGLFDNIAGHLMERQYVRAAPKPSR					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 379>:

a107.seq

```

1  ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGATTG CCGATGCCAA
51  ACTGCCCGTC CTCTTCGACA CCAAGCCCC TACCGGCACC AACTACTGGA
101 TTTATGTGCG CACCGCCCTG CCCGTCTGCC TTGCGTCATT CGGTTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCCTGGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA
351 ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAC
501 CGCCGCGCTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTTCCCCA
551 CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG
601 ACAGGCATCA TCCCGCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCGG
651 CGCAGGCAAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG
701 TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG
751 GAACTCGTCC CCGTATTTAA AGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>:

a107.pep

```

1  MVLTFIWATG GLIADAKLPV LFDTQAPTGT NYWIYVATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151 IADIFKWNDS VSGRTKTAAL TFLPPLISCL LFPTGFVTAI GYVGLAATVW
201 TGIIPAMLLY RSRKKFGAGK TYKVYGLWL MVWVFLFGIV NIAAQVLSQM
251 ELVPVFKG*

```

m107/a107 94.8% identity in 154 aa overlap

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	: : : : :					
a107	MVLTFIWATGGLIADAKLPVLFDTQAPTGTNYWIYVATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60

319

	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
a107	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSFLGVTGLGLFDYIAHLQMERQHLRAAPKPPRX					
a107	TGNMDKILSLFSYMAIATSFLGVTGLGLFDYIADIFKWNDSVSGRTKTAALTFLPPLISCL					
	130	140	150	160	170	180
a107	LFPTGFVTAIGYVGLAATVWTGII PAMLLYRSRKKFGAGKTYKVYGGGLWLMVWVFLFGIV					
	190	200	210	220	230	240

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 381>:

```

g108.seq
1  ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
51  AAcggCGCAT AAAACGCCgc ccTTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATCCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGCG GCAATCTTAC TCGGCGGGTG
201 CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
251 GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TGCGGAATTG
301 GACAGGCGCA GCGAATGGCG TTTGACCGCG CTGGCGATGA GTGCCGAAAA
351 ACAGGCGGAA TGGGAAAACA AGATTTCGCG CTGCGCTACC GAAGAAGCAC
401 CTAACCAGCT GACCGGCAAC GATGTGATGC AGATGCTGaa ccagtccacG
451 CGCaatcagg cacTtgccgc CctgaccgTC AAAacggtTT CcgctgcctT
501 CAaacgcctg tACCGCTAa

```

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

```

g108.pep
1  MLPGFNRIFK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNRKGIE
51  MNKTLILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAEL
101 DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
151 RNQALAAALTV KTVSACFKRL YR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 383>:

```

m108.seq
1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTT CAACACTCGG
51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGA GGCGGTAACA CATTCCGGCAG CTTAGACGGT GGCACAGGCA
251 TGGGCGGCAG CATCGTCAAA ATGGCGGTTG GGAGCCAATG CCGTGCGGAA
301 TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAGTGGGAAA ACAAGATTG CGCTTGCGTC GCCCAAGAAG
401 CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501 CTTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

```

m108.pep
1  MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
51  MNKTLILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVGSQCRAE
101 LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
151 TRNQALAAAL AKTVSACFKH LYR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 108 shows 89.6% identity over a 173 aa overlap with a predicted ORF (ORF 108.ng) from *N. gonorrhoeae*:

m108/g108

320

```

      10      20      30      40      50      60
m108.pep  MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLNILPV
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g108      MLPGFNRIKRFAPTLGTAKHTPPFALSRTGRLIRSYRHKRRGFNRKGIEMNKTLNILPA
          10      20      30      40      50      60

      70      80      90     100     110     120
m108.pep  AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g108      AILLGGCAAGG-NTFGSLDGGTGMGGSIVKMTVESQCRAELDRRSEWRLTALAMSAEKQA
          70      80      90     100     110

      130     140     150     160     170
m108.pep  EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAAALTAKTVSACFKHLYRX
          |||||:|::|:|::|:|::|:|::|:|::|:|::|:|::|:|::|:|::|
g108      EWENKICGCATEEAPNQLTGNDVMQMLNQSTRNQALAAALTAKTVSACFKRLYRX
          120     130     140     150     160     170

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 385>:

a108.seq

```

1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTT CAACACTCGG
51 AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGG GCGGGAACA CATTGCGCAG CTTAGACGGC GGCACAGGTA
251 TGGGCGGCAG CATCGTCAAA ATGGCGGTAG AAAGCCAATG CCGTGGCGAA
301 TTGAACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAATGGGAAA ACAAGATTG CGCTTGGGTC GCCCAAGAAG
401 CACCCAACCA GCTGACCGGC AACGATGTGA TGCAGATGCT GGATCCGCTC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCGCGCTG
501 CTTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 386; ORF 108.a>:

a108.pep

```

1  MLPGFNRIK RFVPTLGTAKHTPPFALSRT GRLIRFYRHK RRGFNKRGIE
51 MNKTLNILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVESQCRAE
101 LNKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPNQLTG NDVMQMLDPS
151 TRNQALAAALT AKTVSACFKH LYS*

```

m108/a108 96.5% identity in 173 aa overlap

```

      10      20      30      40      50      60
m108.pep  MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLNILPV
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a108      MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKTLNILPV
          10      20      30      40      50      60

      70      80      90     100     110     120
m108.pep  AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a108      AILLGGCAAGGGNTFGSLDGGTGMGGSIVKMAVESQCRAELNKRSEWRLTALAMSAEKQA
          70      80      90     100     110     120

      130     140     150     160     170
m108.pep  EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAAALTAKTVSACFKHLYRX
          |||||:|::|:|::|:|::|:|::|:|::|:|::|:|::|:|::|:|::|
a108      EWENKICACVAQEAPNQLTGNDVMQMLDPSTRNQALAAALTAKTVSACFKHLYRX
          130     140     150     160     170

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 387>:

g109.pep
1 MYRRVVGSL DGLGDLAAGI DRRRMLTAFG SGHGNDARQ NHPIRRHRGV
51 LFRLVNPFVG WALTMLLDNL GLIGCKERSA QLG FVGRVLI PAVGFLILCV
101 AMGAVGMLPG IPPFLEQFKS LG

```
m109.seq
1  ATGTATTATC GCCGGGTAT GGGGCTATCC GATGGACTTG GCGATTGGC
51  AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
101 GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
151 CATCGTGGTG TTCTCTCCG CCTTGTC AAT CCGGTTTTCG GCTGGGCGTT
201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAA GAGCGCAGTG
251 CGCAATTAGG TTTCCCGGA GCGCTGTTGA TACCCGCAGT AGGTTTCTTG
301 ATCTTGTTGT TGGCGATGGG TGCGGTCGGG ATGCTGCCCG GTATCCCGCC
351 GTTTTGGGAA CACTTCAAAT CTTTGGGCTA G
```

```
m109.pep
      1  MYYRRVMGLS  DGLGDLAAGI  ERSLGRRRIL  TAFSGHGND  AQRQNHPIRR
     51  HRGVLFRLVN  PVFGWALTML  LDNLGLIGCK  ERSAQLGFAG  RVLIPAVGFL
    101  ILCVAMGAVG  MLPGIPPFLE  HFKSLG*
```

ORF 109 shows 92.9% identity over a 126 aa overlap with a predicted ORF (ORF 109.ng) from *N. gonorrhoeae*:

		10	20	30	40	50	60
m109.pep		MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGND	AQRQNHP	IRRH	RGVLF	RRLVN	
		:					
g109		MYYRRVVGLSDGLGDLAAGIDR	---	RRMLTAFGSGHGND	AQRQNHP	IRRH	RGVLF
		10	20	30	40	50	
		70	80	90	100	110	120
m109.pep		PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLI	PAVGFLIL	CVAMGAVG	MLPGI	PPFLE	
g109		PVFGWALTMLLDNLGLIGCKERSAQLGFVGRVLI	PAVGFLIL	CVAMGAVG	MLPGI	PPFLE	
		60	70	80	90	100	110
m109.pep		HFKSLGX					
		:					
g109		QFKSLGX					
		120					

```

a109.seq
      1  ATGTATTATC  GCCGGGTTGT  GGGGCTATCC  GATGGACTTG  GCGATTGCGC
     51  AGCCGGTATT  GAGCGTAGCC  TTGGTCGTAG  GCGTATACTT  ACCGCTTTTG
    101  GAAGCGGGCA  TGGAAATGAC  GCGCAAAGCG  AAAACCAACC  AATCCGCGCG
    151  CACCGTGGTG  TTCTCTTCGG  CTTGGTCAAT  CCGGTTTTCG  GCTGGGCGTT

```



```

201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGCG
251 CGCAATTAGG TTTCACCGGA CGCGTATTGA TACCCGTAGT AGGTTTCTTG
301 ATCTTGTGTG TGGCGATGGG TCGGGTCGGG ATGCTGCCCG GTATCCCGCC
351 GTTTTTGGAG CACTTCAAAT CTTTGGGCTA G

```

This corresponds to the amino acid sequence <SEQ ID 392; ORF 109>:

a109.pep

```

1 MYYRRVVGLS DGLGLDLAGI ERSIGRRRIL TAFSGHGHND AQRQNHPIRR
51 HRGVLFRVLN PVFGWALTML LDNLGLIGCK ERSAQLGFTG RVLPVVGFL
101 ILCVAMGAVG MLPGIPPFL E HFKSLG*

```

m109/a109 97.6% identity in 126 aa overlap

	10	20	30	40	50	60
m109.pep	MYYRRVMGLSDGLGLDLAGIERSIGRRRILTAFGSGHGHNDAQRQNHPIRRHRGVLFRVLN					
a109	MYYRRVVGLS DGLGLDLAGIERSIGRRRILTAFGSGHGHNDAQRQNHPIRRHRGVLFRVLN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m109.pep	PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLPVVGFLILCVAMGAVGMLPGIPPFL					
a109	PVFGWALTMLLDNLGLIGCKERSAQLGFTGRVLIPVVGFLILCVAMGAVGMLPGIPPFL					
	70	80	90	100	110	120
m109.pep	HFKSLGX					
a109	HFKSLGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 393>:

g111.seq

```

1 ATCCGCTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAaacaacC GCGCAaaccg
101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CcTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGtccaCC TACCAGACCG
251 ATTCCGAAAT CAGCCGGTTt atacagacan atgctggaga gctcttcgcg
301 tntcatgcag nttctataac tgattccgcc gaagactgtc tgcctaatac
351 gcctatctca tcggcgctct ga

```

This corresponds to the amino acid sequence <SEQ ID 394; ORF 111.ng>:

g111.pep

```

1 MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLOGET MGTITYTVKYL
51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMST YQTDSEISRF IQTAGELFAH
101 ASITDSAEDC LPNTPISSAL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 395>:

m111.seq

```

1 ATCCGCTCTG AAACACGCCT GCCGAACCTT ATCCGCGTCT TGATATTTGC
51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101 TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATAYCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AWAACGCAT
201 CGATGACGCG CTTAAAGAAK TCAACCGGyA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTT AACCACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
351 CCGTACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTTCG CCGGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAACAA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA

```

```

601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
701 AGCAGCCCAA TATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
801 TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAATCGC TGTTTTCTG ATTGTCAGGG
1001 ATAAAGGCGG cTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGcTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 396; ORF 111>:

```

m111.pep
  1 MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTITYXVKYL
 51 SNNRDKLPSP AEIXKRIDDA LKEXNRXMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKIIKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPR IGIEQPNIVQ GGNTQTIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

ORF 111 shows 88.7% identity over a 97 aa overlap with a predicted ORF (ORF 111.ng) from *N. gonorrhoeae*:

```

m111.pep/g111.pep

      10      20      30      40      50      60
m111.pep  MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTITYXVKYLSNNRDKLPSP
          |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
g111       MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTITYTVKYLSNNRDKLPSP
          10      20      30      40      50      60

      70      80      90     100     110     120
m111.pep  AEIXKRIDDALKEXNRXMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
          |:| |||:|||:||| || |||| ||||| |||:|:|
g111       AKIQKRIDDALKEVNRQMSTYQTDSEISRFIQTXAGELFAXHAXSITDSAEDCLPNTPIIS
          70      80      90     100     110     120

      130     140     150     160     170     180
m111.pep  GALDVTVGPLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKIIKQKGDYASLSKTHPK
g111       SALX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 397>:

```

a111.seq
  1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCACCT TGATATTTGC
 51 CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101 TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG
351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCGGGAACAA
451 ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAAATCA TTTTGAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAC CCACCCCAAG GCCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA

```

324

```

801 TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAATCTCGC TGTTTTCTTG ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 398; ORF 111.a>:

a111.pep

```

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTITYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMSY YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
151 IKQAASYTGI DKILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPRW IGIEQPNIVQ GGNTQIIIVPL
251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPIHNLAS SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

m111/a111 97.7% identity in 351 aa overlap

m111.pep	10	20	30	40	50	60
	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSP					
a111	MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP					
	10	20	30	40	50	60
m111.pep	70	80	90	100	110	120
	AEIXKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA EAVRLNRLTH					
a111	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA EAVHLNRLTH					
	70	80	90	100	110	120
m111.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASYTGIDKILKQKGDYASLSKTHPK					
a111	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASYTGIDKILKQKGDYASLSKTHPK					
	130	140	150	160	170	180
m111.pep	190	200	210	220	230	240
	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRWIGIEQPNIVQ					
a111	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRWIGIEQPNIVQ					
	190	200	210	220	230	240
m111.pep	250	260	270	280	290	300
	GGNTQIIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNKRPIHNLASISVVADSAM					
a111	GGNTQIIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNKRPIHNLASISVVADSAM					
	250	260	270	280	290	300
m111.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
a111	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 399>:

g111-1.seq

```

1 ATGCCGCTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101 TTACCCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TACCAGACCG

```

325

```

251 ATTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ATTTCCGACA CGTTACCGCC GAAGCCGTCC GCCTGAACCG
351 CCTGACTCAC GCGCGACTGG ACGTAACCGT CGGCCCTTTG GTCAACCTTT
401 GGGGGTTTCG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGCAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAA GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCggcGAGTT
651 GCACGGCAAA GGCAAAAATG CGCACGGCGA ACCGTGGCGC ATCGGTATAG
701 AGCAACCCAA TATcatccaa ggcggcaata cgcAGattat cgtcccgcgtg
751 aaCaaccggt cgttgccac ttccggcgAT taccgtaTT tccacgtcgA
801 TAAAAACGGC Aaacgccttt cccacATCAT CAATCCCAAC AACAAACGAC
851 CCATCAGcCA CAacctcgcc tcCATCagCg TGGTCTCAGA CAGTGCAATG
901 ACGGCGGACG GTTTATCCAC AGGATTATTT GTTTTAGGCG AAACCGAAGC
951 CTTAAGGCTG GCAGAACAAG AAAAActCGC TGTTTTCCCTA ATTGTCCGGG
1001 ATAAGGACGG CTACCGCACC GCCATGTCTT CCGAATTTGC CAAGCTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 400; ORF 111-1.ng>:

g111-1.pep

```

1 MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLOGET MGTTYTVKYL
51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMSY YQTDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKII LQGGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNAGEPWR IGIEQPNIIQ GGNTQIIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVSDSAM
301 TADGLSTGLF VLGETEALRL AEQEKLAFLV IVRDKDGYRT AMSSEFAKLL
351 R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 401>:

m111-1.seq

```

1 ATGCCGTCTG AACACGCCT GCCGAACCTT ATCCGCTCT TGATATTTGC
51 CCTGGGTTTC ATCTTCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101 TTACCTGCA AGCGAAACG ATGGGACGA CCTATACCGT CAAATACCTT
151 TCAATAATC GGGACAAACT CCCCTACCT GCCGAAATAC AAAAACGCAT
201 CGATGACCGG CTTAAAGAAG TCAACCGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAAACAG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
701 AGCAGCCCAA TATCGTCCA GCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACCGT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
801 TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGAGA CAGTGCGATG
901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAGCTG CGAGAGCGG AAAAActCGC TGTTTTCCCT ATTGTGAGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 402; ORF 111-1>:

m111-1.pep

```

1 MPSETRLPNE IRVLIFALGF IFLNACSEQT AQTVTLOGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMSY YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKII LQGGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPR IGIEQPNIVQ GGNTQIIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

m111-1/g111-1 96.6% identity in 351 aa overlap

```

10 20 30 40 50 60
m111-1.pep MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLOGETMGTTYTVKYLSNNRDKLPSP
g111-1 MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLOGETMGTTYTVKYLSNNRDKLPSP

```

sp|P44550|YOJL_HAEIN HYPOTHETICAL LIPOPROTEIN HI0172 PRECURSOR >gi|1074292|pir|IC64144
hypothetical protein HI0172 - Haemophilus influenzae (strain Rd KW20) >gi|1573128 (U32702)
lipoprotein, putative [Haemophilus influenzae Rd] Length = 346
Score = 349 bits (885), Expect = 2e-95
Identities = 177/328 (53%), Positives = 240/328 (72%), Gaps = 4/328 (1%)

Query:	23	LNACSEQTAQTVTLQGETMGTYYVKYLSNNRDKLPSPAEIXKRIDDALKEXNRXMSITYQ	82
		L AC ++T + ++L G+TMGTTY VKYL + S + + I+ LK+ N MSTY+	
Sbjct:	17	LAACQKET-KVISLSGKTMGTYYVKYLDGDSITATS-EKTHEEIEAILKDVNAKMSTYK	74
Query:	83	PDSEISRFNQHT-AGKPLRISSDFAHVTAEAVRLNRLTHGALDVTVGPLVNLWGFPGDKS	141
		DSE+SRFNQ+T P+ IS+DFA V AEA+RLN++T GALDVTVG+P+VNLWGF+P+K	
Sbjct:	75	KDSELSRFNQNTQVNTPIEISADFAKVLAEAIRLNKVTEGALDVTVGPPVNLWGFGEPEKR	134
Query:	142	VTREPSPEQIKQAASYTGIDKIIKQGKDYASLSKTHPKAYLDLSSIAGFGVDKVAGEL	201
		++P+PEQ+ + ++ GIDKI L K+ A+LSK P+ Y+DLSSIAGFGVD+VA +L	
Sbjct:	135	PEKQPTPEQLAERQAWVGIDKITLDTNKEKATLSKALPQVYVDLSSIAGFGVDQVAEKL	194
Query:	202	EKYGIQNYLVEIGGELHGKGNARGEPWRIGIEQPNIVQGGNTQIIVPLNNRSLATSGDY	261
		E+ QNY+VEIGGE+ KGKN G+PW+I IE+P + ++ LNN +A+SGDY	
Sbjct:	195	EQLNAQNYMVEIGGEIRAKGKNIEGKPWQIAIEKPTTGERAVEAVIGLNNMGMASSGDY	254
Query:	262	RIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAMTADGLSTGLFVLGETEALKLA	321
		RI+ ++NGKR +H I+P PI H+LASI+V+A ++MTADGLSTGLFVLGE +AL++A	
Sbjct:	255	RIY-FEENGKRFAHEIDPKTGYPIQHHLASITVLAPTSMTADGLSTGLFVLGEDKALEVA	313
Query:	322	EREKLAVFLIVRDKGGYRTAMSSEFEKL	349
		E+ LAV+LI+R G+ T SS F+KL	
Sbjct:	314	EKNNLAVLYLIIRTDNGFVTKSSSAFKKL	341

a111-1.seq

1	ATGCCGCTCTG	AAACACGCCT	GCCGAACTTT	ATCCGCACCT	TGATATTTCG
51	CCTGAGTTTT	ATCTTCCTGA	ACGCGTGTTC	GGAAACAAAC	GGCGAAACCG
101	TTACCTCTTT	AGGTGAAACG	ATGGGCACGA	CCTATACCTG	CAAAATACCTT
151	TCAAATAATC	GGGACAAACT	CCCTCACCTT	GGCGAAATAC	AAAAGCGCAT
201	CGATCAGCGC	CTTAAAGAGT	TCAACCGGCA	GATGTCCACC	TATCAGCCCG
251	ACTCCGAAAT	CAGCCGGGTC	AACCAACGCA	CAGCGCGCAA	GGCCCTCCCG
301	ATTTCAAGCG	ACTTCGCACA	CGTTACTGCC	GAGCGCGTCC	ACTGAAACCG

```

351 CCTGACACAC GGC GCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAATCA TTTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGA AAAAT ACGGCATTCA AAATTATCTG. GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAACG CGCGCGGCGA ACCTTGCGCG ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
801 TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG
901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTAAAGCTG GCAGAGCGCG AAAA ACTCGC TGTTTTCCTG ATGTGCAGGG
1001 ATAAAGCGG CTACCGCACC GCCATGTCTT CCGAATTGA AAAACTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>:

a111-1.pep

```

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMSY YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKII LKQKDYASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIIVPL
251 NNRLATSGD YRIFHVDKSG KRLSHIINPN NKRPIHNLASISVVDASAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

a111-1/m111-1 98.9% identity in 351 aa overlap

a111-1.pep	10	20	30	40	50	60
	MPSETRLPNFIRTLIFALSFIFLNACSEQT AQTVTLQGETMGTTYTVKYL SNNRDKLPSP					
m111-1	MPSETRLPNFIRVLFALGFIFLNACSEQT AQTVTLQGETMGTTYTVKYL SNNRDKLPSP					
	10	20	30	40	50	60
a111-1.pep	70	80	90	100	110	120
	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA EAVHLNRLTH					
m111-1	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA EAVRLNRLTH					
	70	80	90	100	110	120
a111-1.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKII LKQKDYASLSKTHPK					
m111-1	GALDVTVGPLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKII LKQKDYASLSKTHPK					
	130	140	150	160	170	180
a111-1.pep	190	200	210	220	230	240
	AYLDLSSIAKGFVDKVAGELEKYGIQNYLVEIGGELHGK GKNARGEPWRIGIEQPNIVQ					
m111-1	AYLDLSSIAKGFVDKVAGELEKYGIQNYLVEIGGELHGK GKNARGEPWRIGIEQPNIVQ					
	190	200	210	220	230	240
a111-1.pep	250	260	270	280	290	300
	GGNTQIIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPN NKRPIHNLASISVVDASAM					
m111-1	GGNTQIIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPN NKRPIHNLASISVVDASAM					
	250	260	270	280	290	300
a111-1.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
m111-1	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 405>:

g114.seq

```

1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA
51 GACTTTTTTA TGTCCGCCGG GCGGGACGAG TATGGGCGG TCAATGTCGG

```

```
g114.pep
1  MASITSPLHG AQQECSTFL CPPGGTSMGR SMSVTVGLFC VSINLTISVE
51 YGQSGYFTRA AECKTGCQGI SPSCLNERTV CEVTIKWSSS ETSTSDMACA
101 SRLVNMSSC EGSGEPPGWL CAIIRLSAYS SNASLTISM *
```

```
m114.seq
1  ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCACAGAG AATGCAGCAA
51  GACTTTTTTA TGTCCACCGG GCGGGACGAG TATAGGCGGG TCAATGTCGG
101 TAACGGTAGG TTTGTTTTGT GTTTCACATTA ACTTAACAAT ACTCTGTGAA
151 TACGGTTGAA GCGGCTATTT TATCAGAGCC GCCGCATGTA AAACAGAGTG
201 TCAGGGCATC AACCCGAGCT GTCTGAACGA ACAGACGCTT TGCGAkGTAA
251 CGATAAAATG GTCGAGCAGC GACACATCGA CCAGCGACAT TGCCTGTGCC
301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAArGTTCSg GCGAGCCGcC
351 CgAtAGGTTG TGCAGCAATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401 GTTTGACGAT TTGCGCGATG TAA
```

```
m114.pep
1  MASITSPLHG AHRECSKTFL CPPGGTSIGR SMSVTVGLFC VSINLTISVE
51 YGXSGYFIRA AACKTECQGI NPSCLEQTL CXVTIKWSSS DTSTSDIACA
101 SRLVNMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISM *
```

```

m114/g114 90.0% identity over a 140 aa overlap
              10      20      30      40      50      60
m114.pep    MASITSPLHGAHRECSKTFLCPPGGTSGRSMSTVTVGLFCVSNLTIISVEYGXSGYFIRA
              |||||:::|||||:|||||:|||||:|||||:|||||:|||||
g114        MASITSPLHGAQQECSKTFLCPPGGTSMGRSMSTVTVGLFCVSNLTIISVEYQSGYFTRA
              10      20      30      40      50      60

              70      80      90      100     110     120
m114.pep    AACKTECQGINPSCLNQTLXCXVTIKWSSSDTSTSDIACASRLVNMMSSCEXSGEPPGWL
              ||| |||:|||||:|:| |||||:|||||:|||||:|||||
g114        AECKTGCQGISPSCLNERTVCEVTIKWSSSETSTSDMACASRLVNMMSSCEGSGEPPGWL
              70      80      90      100     110     120

              130     140
m114.pep    CAIIRLSAYSSNASLTISRMX
              |||||:|||||:|||||
g114        CAIIRLSAYSSNASLTISRMX
              130     140

```

```

a114.seq
1  ATGCCGGAGG  CAAGCATCGC  CTCCATCACT  TCGCCGCTGC  ACGGGGCGCA
51  ACAGGAATGC  AGCAAGACTT  TTTTATGTCC  GCCGGGCGGG  ACGAGTATGG
101 GCGCGTCAAT  GTCGGTAACG  GTAGGTTTGT  TTTGTGTTTC  CATTAACTTA
151 ACGATATCTG  TCGAATACGG  TTGAAGCGGC  TATTTTATCA  GAGCCGCCGC
201 ATGTAAACA  GGGTGTCAGG  GCATCAGCCC  GAGCTGCCTG  AACGAACGGA
251 CGGTTTGCGC  CGTTACGATA  AAATGGTCTG  GCAGCGACAC  ATCGACCAGC
301 GACATTGCCT  GTGCCAGCCG  CCTTGTTGAA  ATGATGTCTT  CCTGCGAAGG
351 TTCGGGCGAG  CCGCCCGGAT  GGTGTGTCGC  GATAATCAGG  CTGTCGGCAT
401 ATTCGTCCAA  TCGCCAGTTG  ACAATTTTCA  GGAATGTAA

```

This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>:

a114.pep

```

1 MPEASIASIT SPLHGAQQEC SKTFLCPPGG TSMGRSMSVT VGLFCVSINL
51 TISVEYG*SG YFIRAAACKT GCQGISPSCL NERTVCAVTI KWSSSDTSTS
101 DIACASRLVN MMSSCEGSGE PPGWLCAIIR LSAYSSNASL TISRM*

```

m114/a114 92.9% identity in 140 aa overlap

	10	20	30	40	50
m114.pep	MASITSPLHGAHRECSKTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYGXSG				
a114	MPEASIASITSPLHGAQQECSKTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYGXSG				
	10	20	30	40	50
	60	70	80	90	100
m114.pep	YFIRAAACKTECQGINPSCLNEQTLCXVTIKWSSSDTSTSDIACASRLVNMSSCEXS				
a114	YFIRAAACKTGCGISPSCLNERTVCAVTIKWSSSDTSTSDIACASRLVNMSSCEGS				
	70	80	90	100	110
	120	130	140		
m114.pep	PPGWLCAIIRLSAYSSNASLTISRMX				
a114	PPGWLCAIIRLSAYSSNASLTISRMX				
	130	140			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 411>:

g117.seq

```

1 atggtcgacg aactcgacCT GCTGCCGAT GCCGTCGCCG CCACCCTGCT
51 TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTTCCGAGC
101 GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTGtgga CGAAGTGCAG
151 AAACCTACCC ACTTCGCCCC GGTGGACAGC CTCGCCACGC CGGAAGAAGC
201 CGCACAGCAA GCGGAAACCA TGCAGAAAAT GCTGCTGGCg atggttaccg
251 Acatccgcgt cgtaTTAATC AAACGGCGA TGCGTAcgcg caccCTGcta
301 ttTTtaagCA ACGCCCCCGA CAGCCCTGAA AAACgcgccG TCgccaaAga
351 aacccTCGAC ATCTTCGCCC CGCTCGCCAA CCGCTTGGGC GTGTGGCAGC
401 TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
451 TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
501 ATACATCGAA AACTTCCTCG ATATCTGCG TACGGAATC AAAAAATACA
551 ATATCCAATT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
601 AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGgctTGT TCGACATCCG
651 CGCCGTGCGG ATTCTGGTGC ATACCGTCCC CGaGTGTAC ACCACGCTGG
701 gcaTCGTCCA CAGCCTCTGG CAGCCCATTC CCGGCgagtt CGAcgactAC
751 ATCGCCAACC CCAAAGgcaA CGgttATAAA AGtTTGCACA CCGTCATCGT
801 cggcccGGAA gacaaagggtg tggaaGtgCA AATCCGCACC TTCGAtatGC
851 accAATTCAa CgaatTcggT gtcgccgCCC ACTGGCGtta caaagaaggc
901 ggcaaaaggcg attccGCCTA cgaacaaAAA ATcgccTggt TGCgccaACT
951 CTTGGACTGG CGCGAAAATA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001 CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTtATGT TTTGACCCCG
1051 CACGGCAAAG TCCTCTCTCT GCCAACGGGC GCAACCCCA TCGACTTCGC
1101 CTACGCCCTG CACAGCAGCA TcggCGACCG CTGCCGGGGC GCGAAAGTCG
1151 AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACCG ACAGCGCGTC
1201 GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251 AGGctgGGtc aaATCCGGCA AGGCCATCGG caaAATCCGC GCCTAcatCC
1301 GCCAGcaaaAa cgCgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
1351 AAGCAGCTTG CCAAACCTCAC GCCCAAACCC AACCTGCAAG AGCTTgcccga
1401 aaATCTCGGC taaaaAAAGC cagaagacct ctacacCGCc gtcggacaag
1451 gcgaaatttc caaccgcgcc atCaaaaaag cctgcggcac GCTgaacgaa
1501 ccgccccCGG TGCCCGTCAG CGCAACCACC ATCGTCAAAC AGTCCAAAAT

```



```

1551 CAAAAAAGGT GGCAAAACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601 TGACCACGCT TGCCAAATGC TGCAAAACCG CGCCGCCCGA CGATATTGCC
1651 GGCTTCGTTA CCCGCGAGCG CGGCATTTC GTCCACCGCA AAACCTGCCC
1701 CTCTTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751 GTTGGGCGGC GTTGCAGGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801 CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1851 CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCAGTCC CGCGACTTGG
1901 AAGCCAGCAT GAGGTTACG CTCGAAGTCA AACAAGtCAA CGacCTCCCG
1951 CGCGTCCTCG CCGGCCTCGG CGATGTCAA GCGGTATTGA GCGTTACCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

g117.pep

```

1 MVDELDLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ
51 KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVLLI KLAMRTRTLL
101 FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
151 YREIALLLDE KRTERLEYIE NFDILRTEL KKYNIHFEVA GRPKHIYSIY
201 KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGFDDY
251 IANPKGNGYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG
301 GKGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTP
351 HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
401 EIITAKEGHP SVNWLYEGWV KSGKAIGKIR AYIRQONADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSATT IVKQSKI KKG GKTGVLIDGE DGLMTTLAKC CKPAPPDDIA
551 GFVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTAQTQS RDLEASMRFT LEVKQVNDLP
651 RVLAGLGDVK GVLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 413>:

m117.seq (partial)

```

1 ..GTGAAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCCG GCCGCCCGAA
51 ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
101 ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
151 GAGTGTTACA CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTC
201 CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCTATAAAA
251 GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAAGGCGT GGAAGTACAA
301 ATCCGCACCT TCGATATGCA CCAATTCAAC GAATTCGGTG TCGCCGCCCA
351 CTGgCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCCTAC GAACAGAAAA
401 TCGCCTGGTT GCGCCAATC TTGGACTGGC GCGAAAACAT GGCGGAAAGC
451 GGCAAGGAAG ACCTCGCCGC CGCCTTCAAA ACCGAGCTTT TCAACGACAC
501 GATTTATGTT TTGACCCCGC ACGGCAAGT CCTCTCCCTG CCCACGGGCG
551 CGACCCCAT CGACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCGACCGT
601 TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGCTGT CCACCCGCT
651 CGAAAAACGA CAGCGCGTCG AAATCATTAC CGCCAAAGAA GGGCATCCTT
701 CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
751 AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TCGCGGAAGA
801 AGGCCGCGTC CAACTCGACA AACAGCTTGC CAAACTCACG CCCAAACCCA
851 ACCTGCAAGA GCTTGCCGAA AATCTCGGCT ACAAAGGCC AGAAGACCTC
901 TACACCGCCG TCGGACAAGG CGAAATTTCC AACC CGCCA TCCAAAAAGC
951 CTGCGGCACg CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCACCA
1001 TCGTCAAAACA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
1051 GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
1101 GCCGCCCGAC GATATTATCG GCTTCGTTAC CCGCGAGCGC GgCATTTCAG
1151 TGCACCGCAA AwyyTkCyCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
1201 GAWAAAGTGC TGGACGCAAG CTGGGCGGCA TTGCAGGAAG GACAAGTATT
1251 CGCCGTGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTGTGCGG
1301 ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
1351 ACCCAGTCCC GCGACTTGA AGCCAGCATG AGGTTACGC TCGAAGTCAA
1401 ACAAGTCAAC GACCTCCCGC GCGTCTCGC CAGCCTCGGC GACGTCAAAG
1451 GCGTATTGAG CGTTACCCGG CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

m117.pep (partial)

```

1....VKLKKYNVHF EVAGRPKHIY SIYKKMVKKK LSFDGLFDIR AVRILVDTVP
51  ECYTTLGIVH SLWQPIGGEF DDYIANPKGN GYKSLHTVIV GPEDKGVEVQ
101 IRTFDMHQFN EFGVAHWRY KEGGKGSAY EQKIAWLRQL LDWRENMAES
151 GKEDLAAAFK TELFNDTIYV LTPHGKVL SLPTGATPIDFA YALHSSIGDR
201 CRGAKVEGQI VPLSTPLENG QRVEIITAKE GHPSVNWLYE GWVKSNAIG
251 KIRAYIRQON ADTVREEGRV QLDKQLAKLT PKPNLQELAE NLGYKKPEDL
301 YTAVGQGEIS NRAIQKACGT LNEPPVPVPS ETTIVKQSKI KGGKNGVLI
351 DGEDGLMTTL AKCKPAPPD DIIGFVTRER GISVHRKXXX SFQHLAEHAP
401 XKVLDASWAA LQEGQVFAVD IEIRAQDRSG LLRDVSDALA RHKLNVTAVQ
451 TQSRDLEASM RFTLEVQVN DLPRVLASLG DVKGVLSVTR L*

```

Computer analysis of this amino acid sequence gave the following results:

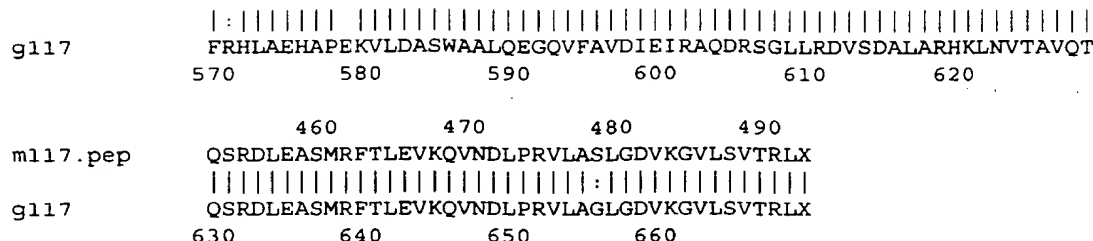
Homology with a predicted ORF from *N. gonorrhoeae*

ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae*:

m117/g117

m117.pep				10	20	30
				VKLKKYNVHF	EVAGRPKHIY	SIYKKMVKKK
g117	EKYREIAL	LLDEKRTER	LEYIENFL	DLIRTEL	KKYNIHFE	VAGRPKHIY
	150	160	170	180	190	200
m117.pep		40	50	60	70	80
		SFDGLFDIR	AVRILVDT	VP	PECYTTLG	IVHSLWQPI
g117		SFDGLFDIR	AVRILVDT	VP	PECYTTLG	IVHSLWQPI
	210	220	230	240	250	260
m117.pep		100	110	120	130	140
		PEDKGVEVQ	IRT	FDMHQF	NEFGVA	AHWRYKEG
g117		PEDKGVEVQ	IRT	FDMHQF	NEFGVA	AHWRYKEG
	270	280	290	300	310	320
m117.pep		160	170	180	190	200
		KEDLAAAFK	TELFNDTI	YVLTPH	GKVL	SLPTGATP
g117		KEDLAAAFK	TELFNDTI	YVLTPH	GKVL	SLPTGATP
	330	340	350	360	370	380
m117.pep		220	230	240	250	260
		PLSTPLENG	QRVEIIT	AKEGHPS	VNWLYE	GWVKSNAI
g117		PLSTPLENG	QRVEIIT	AKEGHPS	VNWLYE	GWVKSNAI
	390	400	410	420	430	440
m117.pep		280	290	300	310	320
		LDKQLAKLT	PKPNLQ	ELAENL	GYKKPE	LDYTAVGQ
g117		LDKQLAKLT	PKPNLQ	ELAENL	GYKKPE	LDYTAVGQ
	450	460	470	480	490	500
m117.pep		340	350	360	370	380
		TTIVKQSKI	KGGKNG	VLI	DGEDGL	MTTLAKC
g117		TTIVKQSKI	KGGKNG	VLI	DGEDGL	MTTLAKC
	510	520	530	540	550	560
m117.pep		400	410	420	430	440
		FOHLAEHAP	XKVL	DASWAA	LQEGQV	FAVDIEI
g117		FOHLAEHAP	XKVL	DASWAA	LQEGQV	FAVDIEI
	570	580	590	600	610	620

332



The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 415>:

a117.seq

```

1  ATGGTTCATG  AACTCGACCT  GCTCCCCGAT  GCCGTCGCCG  CCACCCTGCT
51  TGCCGACATC  GGACGCTACG  TCCCCGACTG  GAACCTATTG  GTTCCGAAC
101 GCTGCAACAG  TACCGTCGCC  GAGCTGGTCA  AAGGTGTGGA  CGAAGTGCAG
151 AAATCAGCCC  ACTTCGCCCC  GGTGGACAGC  CTCGCCACGC  CGGAAGAACG
201 CGCCAGCAG  GCAGAACTA  TGCGGAAAT  GCTGCTGGCG  ATGGTTACCG
251 ACATCCGCGT  CGTGTTAATC  AAATGGCGA  TGCGTACGCG  CACCCTGCAA
301 TTTTAAAGCA  ACGCCCCGA  CAGCCCCGAA  AAACGCGCCG  TCGCAAAGA
351 AACCTCGAC  ATCTCGCCC  CGCTCGCCAA  CCGTTTGGGC  GTGTGGCAGC
401 TCAATGGCA  GTCGAAGAT  TTGGGCTTCC  GCCATCAAGA  ACCCGAAAAA
451 TACCGGAAA  TCGCCCTGCT  TTTGGACGAA  AAACGCACCG  AACGCCTCGA
501 ATACATCGAA  AACTTCCTTA  ATATCCTGCG  TACGGAATC  AAAAAATACA
551 ATATCCACTT  TGAAGTCGCC  GGCCGTCCGA  AACACATCTA  CTCCATTTAC
601 AAAAAATGG  TGAAGAAAAA  ACTCAGCTTC  GACGGGTGTG  TCGACATCCG
651 CGCCGTGCGG  ATTCTGGTTG  ATACCGTCCC  CGAGTGTTAC  ACCACACTGG
701 GCATTGTCCA  CAGCCTCTGG  CAGCCCATTC  CCGGCGAGTT  CGACGACTAC
751 ATCGCCAACC  CGAAAGGCAA  CGGCTATAAA  AGTTTGCACA  CCGTCATCGT
801 CGGCCCGGAA  GACAAAGGCG  TGGAAGTGCA  AATCCGCACC  TTCGATATGC
851 ACCAATTCAA  CGAATTCGGT  GTCGCCGCGC  ACTGGCGTTA  CAAAGAGGGC
901 GGCAAAGGCG  ATTCCGCCTA  CGAACAAAAA  ATCGCCTGGT  TACGCCAACT
951 TTTGGACTGG  CGCGAAAACA  TGGCGGAAAG  CGGCAAGGAA  GACCTCGCCG
1001 CCGCCTTCAA  AACCAGAGCT  TTCAACGACA  CGATTTATGT  TTTGACCCCG
1051 CACGGCAAAG  TCCTCTCCCT  GCCCACAGGC  GCGACCCCA  TCGACTTCGC
1101 CTACGCCCTG  CACAGCAGCA  TCGGCGACCG  TTGCCGCGGT  GCGAAAGTCG
1151 AAGGGCAGAT  TGTGCCGCTG  TCCACCCCGC  TCGAAAACGG  ACAGCGTGTC
1201 GAAATCATTA  CCGCCAAAGA  AGGGCATCCT  TCCGTCAACT  GGCTTTACGA
1251 AGGCTGGGTC  AAATCCAACA  AGGCAATCGG  CAAAATCCGC  GCCTACATCG
1301 CCCAGCAAAA  CGCGACACC  GTGCGCGAAG  AAGGCCGCGT  CCAACTCGAC
1351 AAACAGCTTG  CCAAACCTAC  GCCCAAACCC  AACCTGCAAG  AGCTTGCCGA
1401 AAATCTCGGC  TACAAAAAGC  CAGAAGACCT  CTACACCGCC  GTCGGACAAG
1451 GCGAAATTTT  CAACCGCGCC  ATCCAAAAAG  CCTGCGGCAC  GCTGAACGAA
1501 CCGCCGCCCG  TACCGTCCAG  CGAAACCACC  ATCGTCAAAC  AGTCCAAAAT
1551 CAAAAAAGGC  GGCAAAAACG  GCGTGCTCAT  CGACGGCGAA  GACGGTCTGA
1601 TGACCACGCT  TGCCAAATGC  TGCAAACCCG  CGCCGCCCGA  CGACATTGTC
1651 GGCTTCGTTA  CCCGCGATCG  CGGCATTTTC  GTACACCGCA  AAACCTGCCC
1701 CTCTTTCCGA  CACCTCGCCG  AACACGCGCC  CGAAAAAGTA  CTGGACGCAA
1751 GTTGGGCGGC  GTTGCAGGAA  GGACAAGTGT  TCGCCGTCGA  TATCGAAATC
1801 CGCGCCCAAG  ACCGCTCCGG  GCTTTTGCGC  GACGTATCCG  ACGCGCTCGC
1851 CCGCCACAAA  CTCAACGTTA  CCGCCGTGCA  AACCCAGTCC  CGCGACTTGG
1901 AAGCCAGCAT  GAGGTTACAG  CTCGAAGTCA  AACAAGTTAC  CGACCTCCCA
1951 CGCGTCTCG  CCAGCCTCGG  CGACGTCAAA  GGCGTATTGA  GCGTTACCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 416; ORF 117.a>:

a117.pep

```

1  MVHELDLLPD  AVAATLLADI  GRYVPDWNLL  VSERCNSTVA  ELVKGVDEVQ
51  KLTHFARVDS  LATPEERAQQ  AETMRKMLLA  MVTDIRVLI  KLAMRTRTLQ
101 FLSNAPDSPE  KRAVAKETLD  IFAPLANRLG  VWQLKWQLED  LGFRHQEPEK
151 YREIALLLDE  KRTERLEYIE  NFLNLRTEL  KKYNIHFEVA  GRPKHIYSIY
201 KKMVKKKLSF  DGLFDIRAVR  ILVDTVPECY  TTLGIVHSLW  QPIPGFDDY
251 IANPKNGYK  SLHTVIVGPE  DKGVEVQIRT  FDMHQFNEFG  VAAHWRYKEG
301 GKGDSAYEQK  IAWLRQLLDW  RENMAESGKE  DLAAAFKTEL  FNDTIYVLTP
351 HGKVLSLPTG  ATPIDFAYAL  HSSIGDRCRG  AKVEGQIVPL  STPLENGQRV

```

333

401 EIITAKEGHP SVNWLYEGWV KSNKAIGKIR AYIRQQNADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSETT IVKQSKIKKG GKNGLVIDGE DGLMTTLAKC CKPAPPDDIV
551 GFVTRDRGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GOVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTAQTQS RDLEASMRFT LEVKQVTDLP
651 RVLASLGDVK GVLSVTRL*

m117/a117 98.0% identity in 490 aa overlap

m117.pep				10	20	30
				VKLKKYNVHFEVAGRPKHIYSIYKKMVKKKL		
a117	EKYREIALLLDEKRT	LEYIENFLN	ILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKKL			
	150	160	170	180	190	200
m117.pep		40	50	60	70	80
	SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVG					
a117	SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVG					
	210	220	230	240	250	260
m117.pep		100	110	120	130	140
	PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESG					
a117	PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESG					
	270	280	290	300	310	320
m117.pep		160	170	180	190	200
	KEDLAAAFKTELFNDTIYVLT	PHGKVL	SLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV			
a117	KEDLAAAFKTELFNDTIYVLT	PHGKVL	SLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV			
	330	340	350	360	370	380
m117.pep		220	230	240	250	260
	PLSTPLENGQORVEIITAKEGHPSVNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQ					
a117	PLSTPLENGQORVEIITAKEGHPSVNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQ					
	390	400	410	420	430	440
m117.pep		280	290	300	310	320
	LDKQLAKLTPKPNLQELAENLG	YKKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSE				
a117	LDKQLAKLTPKPNLQELAENLG	YKKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSE				
	450	460	470	480	490	500
m117.pep		340	350	360	370	380
	TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKXXXS					
a117	TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIVGFVTRDRGISVHRKTCPS					
	510	520	530	540	550	560
m117.pep		400	410	420	430	440
	FQHLAEHAPXKVL	DASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT				
a117	FRHLAEHAPEKVL	DASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT				
	570	580	590	600	610	620
m117.pep		460	470	480	490	
	QSRDLEASMRFTLEVQVNDLPRVLASLGDVKGVLSVTRLX					
a117	QSRDLEASMRFTLEVQVNDLPRVLASLGDVKGVLSVTRLX					
	630	640	650	660		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 417>:

g117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
101 AAAACCTCAT CGGTACCGCA TGGTCGCTGG CGCAGGAACA TTATCCTGCC
151 GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201 GGCAGCAATG GTGACGGAAC TCACCTGCTG GCCCGATGCC GTCGCCGCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT
301 TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351 AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401 AAGAACGCGC ACAGCAAGCG GAAACCATGC GGAAAATGCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCTGAAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG ATTGGGCGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTG GCGTTCCGCC ATCAAGAACC
651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAAC TTCTCGATA TCCTGCGTAC GGAACTCAAA
751 AAATACAATA TCCACTTTGA AGTCGCCGGC GTCCGAAAC ACATCTACTC
801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCTGTTCG
851 ACATCCGCGC CGTGC GGATT CTGGTCGATA CCGTCCCCGA GTGTACACC
901 ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagttCGA
951 cgactACATC GCCAACCCCA AAGgcaACGg ttATAAAAGt TTGCACACCG
1001 TCATCGTcgg cccGGAagaa aaagggtgtg aagtgcAAAT CCGCACCTTC
1051 GATATGCacc AATTcCaaCga ATTCCGGTGT CCGCGCCACT GCGGTTACAA
1101 AGAAGGCGGC AAAGGCGATT CCGCTACGA ACAAAAATC GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAATATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCGC CCTTCAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG
1301 ACTTCGCTTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG
1351 AAAGTCGAAG GGCAGATTGT GCGCTGTCC ACCCGCTCG AAAACGGACA
1401 CGCGCTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAA TCCGGCAAGG CCATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGTGTCCA
1551 ACTCGACAAG CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 TGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT CGGCGACGCT
1701 GAACGAACCG CCGCCCGTGC CCGTCAGCGC AACCAACATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGGCGAAGAC
1801 GCGTTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 TATTGCCGGC TTCGTTACCC GCGAGCGCGG CATTTCGCTC CACCGCAAAA
1901 CCTGCCCTTC TTTCCGACAC CTGCGCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTCG CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCGCGC GTCTCGCCG GCCTCGGCGA TGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTA

```

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

g117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WSLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VDELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK
251 KYNHFVEVAG RPKHIYSIYK KMKVKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKGNKYKS LHTVIVGPPE KGVVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTPE KVLVSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQVVE IITAKEGHPN VNWLYEGWVK SGKAIGKIRA
501 YLRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNPE PPVPVSATTI VKQSKIKKGG KTGVLIDGED
601 GLMTTLAKCC KPAPPDDIAG FVTRERGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 419>:

m117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA

```

```

51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCGG GACAACGATA
101  AAAACCTCAT CGGTACCGCA TGGTTGCTGG CGCAGGAACA TTACCCCGCC
151  GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201  GGCGCAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCGGCCA
251  CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301  TCCGAACGCT GCAACAGTAC CGTCGCGGAG CTGGTCAAAG GTGTGGACGA
351  AGTGCAGAAA CTCACCCACT TCGCCCCGGT GGACAGCCTC GCCACGCCGG
401  AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAATGCT GCTGGCGATG
451  GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501  CCTGCAATTT TTAAGCAACG CCCCAGACAG CCCCAGAAAA CGCGCCGTCG
551  CCAAAGAAAC CCTCGACATC TTCGCCCGC TCGCCAACCG TTTGGGCGTG
601  TGGCAGCTCA AATGGCAGCT CGAAGATTG GGCTTCCGCC ATCAAAAGCC
651  CGAAAAATAC CGCGAAATCG CGCTGCTTTT GGACGAAAAA CGCACCGAAC
701  GCCTCGAATA CATCGAAAAC TTCCTCAACA TCCTGCGCGG TGAACTCAAG
751  AAATACAATG TCCATTTCTGA AGTCGCGCGC CGCCCGAAAC ACATCTACTC
801  CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTCTTTG
851  ACATCCGCGC CGTGCGAATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
901  ACGCTGGGTA TCGTCCACAG CCTCTGGCAG CCATTCCCG GCGAGTTCGA
951  CGACTACATC GCCAATCCCA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTACAAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCCCACT GGCGTTACAA
1101 AGAGGGCGGC AAGGGCGATT CCGCCTACGA ACAGAAAAAT GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCCGCC CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGGAC GCGAAAGTCC TCTCCCTGCC CACGGGCGCG ACCCCCATCG
1301 ACTTCGCTTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCAGGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGTAC AAAAAGCCAG AAGACCTCTA CACCGCGCTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAATCAA AAAAGCGGC AAAACGCGC TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCGCGC CGCCGACGA
1851 TATTATCGGC TTCGTTACCC GCGAGCGCGG CATTTAGTG CACCGCAAAA
1901 CCTGCCCGTC TTTCCAACAC CTCGCCGAAC ACGCGCCCGA AAAAGTGTGT
1951 GACGCAAGCT GGGCGGCATT GCAGGAAGGA CAAGTATTCT CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GTCGCGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAATC AACGTTACCG CCGTGCAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCGCGC GTCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTA

```

This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>:

ml17-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WLLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101  SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151  VDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201  WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNILRGELK
251  KYNVHFVEVAG RPKHIYSIYK KMKVKKLSFD GLFDIRAVRI LVDTVPECYT
301  TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351  DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLDWR ENMAESGKED
401  LAAAFKTELF NDIYVLTTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451  KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501  YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551  GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601  GLMTTLAKCC KPAPPDDIIG FVTRERGISV HRKTCPSFQH LAEHAPEKVL
651  DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701  DLEASMRFTL EVKQVNDLPR VLASLGDVKG VLSVTRL*

```

ml17-1/g117-1 98.2% identity in 737 aa overlap

	10	20	30	40	50	60
ml17-1.pep	MTAISPIQDTQSATLQELREWFDSDYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
g117-1	MTAISPIQDTQSATLQELREWFDSDYCAALPDNDKNLIGTAWSLAQEHYPADAATPYGEPL					
	10	20	30	40	50	60
	70	80	90	100	110	120

m117-1.pep PDHFLGAAQMVHEDLLPDVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
|||||
g117-1 PDHFLGAAQMVDELDDLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
70 80 90 100 110 120

m117-1.pep 130 140 150 160 170 180
LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK
|||||
g117-1 LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK
130 140 150 160 170 180

m117-1.pep 190 200 210 220 230 240
RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN
|||||
g117-1 RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN
190 200 210 220 230 240

m117-1.pep 250 260 270 280 290 300
FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
|||||
g117-1 FLDILRTELKKYNHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
250 260 270 280 290 300

m117-1.pep 310 320 330 340 350 360
TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
|||||
g117-1 TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEEKGVEVQIRTFDMHQFNEFGV
310 320 330 340 350 360

m117-1.pep 370 380 390 400 410 420
AAHWRYKEGGKGSAYEQKIAWLRLQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH
|||||
g117-1 AAHWRYKEGGKGSAYEQKIAWLRLQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH
370 380 390 400 410 420

m117-1.pep 430 440 450 460 470 480
GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
|||||
g117-1 GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
430 440 450 460 470 480

m117-1.pep 490 500 510 520 530 540
VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
|||||
g117-1 VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
490 500 510 520 530 540

m117-1.pep 550 560 570 580 590 600
KKPEDLYTAVGQGEISNRAIQKACGTLENEPPVPVSETTIVKQSKIKKGGKNGVLIDGED
|||||
g117-1 KKPEDLYTAVGQGEISNRAIQKACGTLENEPPVPVSATTIVKQSKIKKGGKNGVLIDGED
550 560 570 580 590 600

m117-1.pep 610 620 630 640 650 660
GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG
|||||
g117-1 GLMTTLAKCKPAPPDDIAGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDASWAALQEG
610 620 630 640 650 660

m117-1.pep 670 680 690 700 710 720
QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVAVQTQSRDLEASMRFTLEVQVNDLPR
|||||
g117-1 QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVAVQTQSRDLEASMRFTLEVQVNDLPR
670 680 690 700 710 720

m117-1.pep 730
VLASLGDVKGVLVSVTRLX
|||||
g117-1 VLAGLGDVKGVLVSVTRLX
730

m117-1/RelA

sp|P55133|RELA_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP SYNTHETASE I) >gi1537617 (U13769) ppGpp synthetase I [Vibrio sp.] Length = 744
 Score = 536 bits (1366), Expect = e-151
 Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)

Query: 74 LDLLPDAVAATLLADI---GRYVPDWNLLVSERCNSTVAELVKGVDVQKLTHFARVDSL 130
 L + D + A LL + G Y D + E + T+ LV+GV+++ ++ ++ S
 Sbjct: 68 LSMADATLIAALLYPLVEGGCYSTD---ALKEEYSGTILHLVQGVQMC AIS---QLKST 121

Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLKIAMRTRTLQFLSNAPDSPEKRAVAKETLDI 190
 A +A Q + +R+MLL+MV D R V+IKLA R L+ + + PD +RA A+E +I
 Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFCVVIKLAERICNLREVKDQPDDEV-RRAAAQECANI 180

Query: 191 FAPLANRLGVWQLKWQLEDLGRFHQKPEKYREIALLLDEKRTERLEYIENFLNIRGELK 250
 +APLANRLG+ QLKW++ED FR+Q P+ Y++IA L E+R +R +YI +F++ L +K
 Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDQTYKQIAQLSERRIDREDYITHFVDDLSAMK 240

Query: 251 KYNVHFEVAGRPKHIYSIYKMKVKKKLSFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQ 310
 N+ EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY LG+VH+ ++
 Sbjct: 241 ASNIRAEVQGRPKHIYSIWRKMQKKSLEFDELFDVRAVRIVAEELQDCYAALGVVHTKYR 300

Query: 311 PIPGEFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRT FDMHQFNEFGVAAHWRYKEG- 369
 +P EFDDY+ANPK NGY+S+HTV++GPE K +E+QIRT MH+ +E GVAHW+YKEG
 Sbjct: 301 HLPKEFDDYVANPKNGYQSIHTVVLGPEGKTIEIQIRTKQMHEESELGVAHWKYKEGT 360

Query: 370 --GKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLT PHGKVL SLP 427
 G SAY++KI WLR+LL W+E M++SG ++ +++++F+D +Y TP G V+ LP
 Sbjct: 361 ASGGAQSAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFD RRVYAFTPKGDVVDLP 418

Query: 428 TGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPSVNWLYE- 486
 + ATP+DFAY +HS +G RC GAKVEG+IVP + L+ G +VEIIT KE +PS +WL
 Sbjct: 419 SNATPLDFAYHIHSEVGHRCIGAKVEGRIVPFTYHLQMGDQVEIITQKEPNPSRDWLNPN 478

Query: 487 -GWVKSNAKIGKIRAYIRQONADTVREEGRVQLDKQLAKL--TPKPNLQELAENLG YKKP 543
 G+V S++A K+ A+ R+Q+ D G+ L+ +L K+ T K A+ K P
 Sbjct: 479 LGFVTSSRAKVAHWFRKQDRDKNIIAGKEILEAELVKIHA TLKDAQYAAKRFNVKSP 538

Query: 544 EDLYTAVGQGEIS-NRAIQKACGTLNEPPVPVSETTIVKQSKI-----KKGKNGV 594
 E+LY +G G++ N+ I +N+P + + K S+ KK ++ V
 Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEDQQLLEKLSEASNKQATSHKKPQRDAV 598

Query: 595 LIDGEDGLMTTLAKCKCPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDA SW 654
 +++G D LMT LA+CC+P P DDI GFVT+ RGISVHR C + L HAPE+++D W
 Sbjct: 599 VVEGVDNLMTHLARCCQPIPGDDIQGFVTQGRGISVHRMDCEQLEELRH HAPERIIDTVW 658

Query: 655 AALQEGQVFAVDIEIRAQDRSGLLRDVS DALARHKLNVTA VQTQ--SRDLEASMRFTLEV 712
 G + + + + A +R+GLL++++ L K+ V ++++ + + M F LE+
 Sbjct: 659 GGGFVGN-YTITVRVTASERNGLLKELTNTLMNEKVKVAGMKSRVDYKKQMSIMDFE LEL 717

Query: 713 KQVNDLPRVLASLG DVKGVLSVTRL 737
 + L RVL + VK V RL
 Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEAKRL 742

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 421>:

al17-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCAC CGCGCTGCCG AACAAACGATA
101  AAAAAGTTGT CTTAGCCGCC CGTTCGCTGG CGGAAGCACA TTACCCCGCC
151  GATGCCGCCA CGCCGTATGG CGAACCCTG CCGGACCACT TCCTCGGCGC
201  GGCGCAATG GTTCATGAAC TCGACCTGCT CCCCAGTACC GTCGCGGCCA
251  CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301  TCCGAACGCT GCAACAGTAC CGTCGCGGAG CTGGTCAAAG GTGTGGACGA
351  AGTGCAGAAA CTCACCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401  AAGAAGCGCG CCAGCAGGCA GAAACTATGC GGAAATGCT GCTGGCGATG
451  GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501  CCTGCAATTT TTAAGCAACG CCCCAGACAG CCCCAGAAAA CGCGCCGTCG
551  CCAAGAAAC CCTCGACATC TTCGCCCCGC TCGCAACCG TTTGGGCGTG
601  TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTCCGCC ATCAAGAACC
651  CGAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701  GCCTCGAATA CATCGAAAAA TTCCTTAATA TCCTGCGTAC GGAAC TCAA
751  AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
801  CATTACAAA AAAATGCTGA AGAAAAAACT CAGCTTCGAC GGGTTGTTCC

```



```

851 ACATCCGCGC CGTGCGGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
901 ACACTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTGCAAAT CCGCACCTTC
1051 GATATGCACC AATTCACGA ATTCGGTGTC GCCGCGCACT GGCGTTACAA
1101 AGAGGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAAIC GCCTGGTTAC
1151 GCCAACTTTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCGC CTTTCAAAC CGAGCTTTTC AACGACACGA TTTATGTTT
1251 GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGTGTTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCG AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AATTTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 CATTGTTCGG TTGCTTACCC GCGATCGCGG CATTTCCGTA CACCGCAAAA
1901 CTTGCCCTCT TTCCGACAC CTCGCCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGA CAAGTGTTCC CCGTCGATAT
2001 CGAAATCCCG GCCCAAGACC GTCGCGGCT TTTGCGGAC GTATCCGACG
2051 CGCTCGCCCG CCACAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTACGCTC GAAGTCAAAC AAGTTACCGA
2151 CCTCCCACGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTAA

```

This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

a117-1.pep

```

1 MTAISPIQDT QSATLQELRE WFDSYCTALP NNDKKLVLA RSLAEAHYPA
51 DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNLRTELK
251 KYNIHFEVAG RPKHIYSIYK KMKVKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVTDLFR VLASLGDVKG VLSVTRL*

```

a117-1/m117-1 97.7% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFDSSYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
a117-1	MTAISPIQDTQSATLQELREWFDSSYCTALPNNDKKLVLAARSRLAEAHYPADAATPYGEPL					
	10	20	30	40	50	60
m117-1.pep	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
a117-1	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
	70	80	90	100	110	120
m117-1.pep	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK					
a117-1	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEK					
	130	140	150	160	170	180
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFHQKPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240

m117-1.pep	250	260	270	280	290	300
	FLNILRGELKKYNVHFVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT					
a117-1	FLNILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT					
	250	260	270	280	290	300
m117-1.pep	310	320	330	340	350	360
	TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTQFDMHQFNEFGV					
a117-1	TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTQFDMHQFNEFGV					
	310	320	330	340	350	360
m117-1.pep	370	380	390	400	410	420
	AAHWRYKEGGKGSAYEQKIAWLRLQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
a117-1	AAHWRYKEGGKGSAYEQKIAWLRLQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
	370	380	390	400	410	420
m117-1.pep	430	440	450	460	470	480
	GKVLSTLPTGATPIDFAYALHSSIGDRRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS					
a117-1	GKVLSTLPTGATPIDFAYALHSSIGDRRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS					
	430	440	450	460	470	480
m117-1.pep	490	500	510	520	530	540
	VNWLYEGWVKSNAIGKIRAYIRQONADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
a117-1	VNWLYEGWVKSNAIGKIRAYIRQONADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
	490	500	510	520	530	540
m117-1.pep	550	560	570	580	590	600
	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKGGKNGVLIDGED					
a117-1	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKGGKNGVLIDGED					
	550	560	570	580	590	600
m117-1.pep	610	620	630	640	650	660
	GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDAASWALQEG					
a117-1	GLMTTLAKCKPAPPDDIIGFVTRDRGISVHRKTCPSFRHLAEHAPEKVLDAASWALQEG					
	610	620	630	640	650	660
m117-1.pep	670	680	690	700	710	720
	QVFAVDIEIRAQDRSGLLRDVS DALARHKLNTAVQTQSRDLEASMRFTLEVQVNDLPR					
a117-1	QVFAVDIEIRAQDRSGLLRDVS DALARHKLNTAVQTQSRDLEASMRFTLEVQVNDLPR					
	670	680	690	700	710	720
m117-1.pep	730					
	VLASLGDVKGVLSTRLX					
a117-1	VLASLGDVKGVLSTRLX					
	730					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 423>:

```

g118.seq
1  ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATcgaGGT TAGGAGAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAATTG GGAGCTTTTT GAAATTAAAG
251 CTTCCCTTG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAaCa AAAAATAG

```

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

```

g118.pep
1  MCEFKDFRRN IPCFEEYDEN SFIGKWYDDG VWDDEEYWKL ENDLIEVRRK
51  YPYPMIDIPRD IVIGIGTIID FLMVFNWELF EIKASFWLPD SVGIHERYER
101 FTTMLRYIFT EKDIVNVRFD YNNKK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 425>:

```
m118.seq
1  ATGTGTGAGT TCAAGGATAT TATAAGAAAC GTTCCTTATT TTGAGGGGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGACAT ACCAAGATAT GTTGTCAATTG GAATCGGTAC
201 CATTATTGAT TTCTTAATGG TTCCAAATTG GAAACTTTTT GAAATTAAAG
251 CTTCCCTTG GTTGCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 426; ORF 118>:

```
m118.pep
1  MCEFkdIIRN VPyfEGyDEN sFIGkwyDDG VWDDEEYwKL ENdLIEVRkK
51  YPyPMDIPRY VVIGIGTIID fLMVPnWkLf EIKASpWLPd SVGIHERYER
101 FTTMLRYIFT EKdIVNVrFD YYNKK*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 118 shows 92.8% identity over a 125 aa overlap with a predicted ORF (ORF 118.ng) from *N. gonorrhoeae*:

```
m118/g118

          10          20          30          40          50          60
m118.pep  MCEFkdIIRNVPyfEGyDENsFIGkwyDDGVWDDEEYwKLEndLIEVRkKYPyPMDIPRY
          |||||: ||:| || |||||:|||||:|||||:|||||:|||||:|||||:|||||:
g118       MCEFkdFRRNIPCFEeYDENsFIGkwyDDGVWDDEEYwKLEndLIEVRkKYPyPMDIPRD
          10          20          30          40          50          60

          70          80          90          100         110         120
m118.pep  VVIGIGTIIDfLMVPnWkLfEIKASpWLPdSVGIHERYERFTTMLRYIFtEKdIVNVrFD
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g118       IVIGIGTIIDfLMVPnWELfEIKASpWLPdSVGIHERYERFTTMLRYIFtEKdIVNVrFD
          70          80          90          100         110         120

m118.pep  YYNKKX
          |||||
g118       YYNKKX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 427>:

a118.seq

```
1  ATGTGTGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAATTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAAATTG GGAGCTTTTT GAAATTAAAG
251 CTTCCCTTG GTTGCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 428; ORF 118.a>:

a118.pep

```
1  MCEFkdFRRN IPCFEeYDEN sFIGkwyDDG VWDDEEYwKL ENdLIEVRkK
51  YPyPMDIPRD IVIGIGTIID fLMVPnWELf EIKASpWLPd SVGIHERYER
101 FTTMLRYIFT EKdIVNVrFD YYNKK*
```

m118/a118 93.6% identity in 125 aa overlap

```
          10          20          30          40          50          60
m118.pep  MCEFkdIIRNVPyfEGyDENsFIGkwyDDGVWDDEEYwKLEndLIEVRkKYPyPMDIPRY
          |||||: ||:| || |||||:|||||:|||||:|||||:|||||:|||||:|||||:
a118       MCEFkdFRRNIPCFEeYDENsFIGkwyDDGVWDDEEYwKLEndLIEVRkKYPyPMDIPRD
          10          20          30          40          50          60
```

		70	80	90	100	110	120
m118.pep		VVIGIGTIIDFLMVPNWKLFEIKAS	PWLPDSVGIHERYERFTT	MLRYIFTEKDIVNVRFD			
		: : : : : : : :					
a118		IVIGIGTIIDFLMVPNWELFEIKAS	PWLPDSVGIHERYERFTT	MLRYIFTEKDIVNVRFD			
		70	80	90	100	110	120
m118.pep	YYNKKX						
a118	YYNKKX						

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 429>:

```
g120.seq
1  ATGATGAAGA  CTTTTAAAAA  TATATTTTCC  GCCGCCATTT  TGTCCGCCGC
51  CCTGCCGTGC  GCGTGTGCGG  CAAGGCTACC  CCAATCCGCC  GTGCTGCACT
101  ATTCGGCGAG  CTACGGCATT  CCCGCCACGA  TGACATTTTA  ACGCAGCGGC
151  AATGCTTACA  AAATCGTTTC  GACGATTAAA  GTGCCGCTAT  ACAATATCCG
201  TTTCGAATCC  GGCGGTACGG  TTGTCGGCAA  TACCCTGCAC  CCTGCCTACT
251  ATAAAGACAT  ACGCAGGGGC  AAACGTGTAT  CGGAAGCCAA  ATTCGCCGAC
301  GGCAGCGTAA  CCTACGGCAA  AGCGGGCGAG  AGCAAAACCG  AGCAAAGCCC
351  CAAGGCTATG  GATTTGTTCA  CGCTTGCTGT  GCAGTTGGCG  GCAAATGACG
401  CGAAACTCCC  CCCGGTCTGT  AAAATCACA  ACGGCAAAAA  GCTTTATTCC
451  GTCGGCGGCC  TGAATAAGGC  GGGTACGGGA  AAATACAGCA  Taggcggcgt
501  gGAAACCGAA  GTCGTCAAAT  ATCGGGTGCG  GCGCGGCGAC  GATACGGTAA
551  CGTATTTCTT  CGCACCGTCC  CTGAACAATA  TTCCGGCACA  AATCGGCTAT
601  ACCGAcgaCG  GCAAAACCTA  TACGTGAAG  CTCAAATCGG  TGCAGATCAA
651  CGGACAGGCC  GCCAAACCGT  AA
```

This corresponds to the amino acid sequence <SEQ ID 430; ORF 120.ng>:

```

g120.pep
  1  MMKTFKNIFS AAILSAAALPC AYAARLPQSA VLHYSGSYGI PATMTFERSG
 51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PAYYKDIRRG KLYAEAKFAD
101  GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151  VGGLNKAGTG KYSIGVGETE VVKYRVRRGD DTVTYFFAPS LNNIPAQIGY
201  TDDGKTYTLK LKSVOINGOA AKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 431>:

```

m120.seq
1  ATGATGAAGA  CTTTTAAAAA  TATATTTTCC  GCCGCCATTT  TGTCCGCCGC
51  CCTGCCGTGC  GCGTATGCGG  CAGGGGTGCC  CCAATCCGCC  GTGCTGmACT
101 AATCCGGCAG  CTACGGCATT  CCCGCCACGA  TGACATTTTA  ACGCAGCGCG
151 AATGCTTACA  AAATCGTTTC  GACGATTAAA  GTGCCGTAT  ACAAATATCCG
201 TTTGAGTCC  GCGGTACGG  TTGTCGGCAA  TACCCTGCAC  CCTACCTACT
251 ATAGAGACAT  ACGCAGGGGC  AAACGTGTATG  CGGAAGCcAA  ATTCGCCGAC
301 GGCAGCGTAA  CTTACGGCAA  AGCGGGCGAG  AGCAAAACCG  AGCAAAGCCC
351 CAAGGCTATG  GATTTGTTCA  CGCTTGCCCTG  GCAGTTGGCG  GCAAATGACG
401 CGAAACTCCC  CCCGGGGCTG  AAAATACCA  ACGGCAAAAA  ACTTTATTCC
451 GTCGGCGGTT  TGAATAAGGC  GGGTACAGGA  AAATACAGCA  TAGGCGGCGT
501 GGAAACCGAA  GTCGTCAAAT  ATCGGTTGCG  GCGCGGCGAC  GATGCGGTAA
551 TGTATTTCTT  CGCACCCTCC  CTGAAACAATA  TTCCGGCACA  AATCGGCTAT
601 ACCGACGACG  GCAAAACCTA  TACGTGAAA  CTCAAATCGG  TGCAGATCAA
651 CGGCCAGGCA  GCCAAACCG

```

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>:

m120.pep

1	<u>MMKTFKNIFS</u>	<u>AAILSALPC</u>	AYAAGLPQSA	VLXYSGSYGI	PATMTFERSG
51	NAYKIVSTIK	VPLYNIRFES	GGTVVGNTLH	PTYRDIRRG	KLYAEAKFAD
101	GSVTY GKAGE	SKTEQSPKAM	DLFTLAWQLA	ANDAKLPPGL	KITNGKKLYS
151	VGGLNKAGTG	KYSIGGVETE	VVKYRVRRGD	DAVMYFFAPS	LNNIPAQIGY
201	TDDGKTYTLK	LKSQVINGOA	AKP		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 120 shows 97.3% identity over a 223 aa overlap with a predicted ORF (ORF 120.ng) from *N. gonorrhoeae*:

m120/g120

	10	20	30	40	50	60
m120.pep	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLXYSYGI PATMTFERSGNAYKIVSTIK					
g120	MMKTFKNIFSAAILSAALPCAYAARLPQSAVLHYSGSYGI PATMTFERSGNAYKIVSTIK					
	10	20	30	40	50	60
	70	80	90	100	110	120
m120.pep	VPLYNIRFESGGTVVGNLHPTYYRDIRRGKLYAEAKFADGSVTY GKAGESKTEQSPKAM					
g120	VPLYNIRFESGGTVVGNLHPAYYKDIRRGKLYAEAKFADGSVTY GKAGESKTEQSPKAM					
	70	80	90	100	110	120
	130	140	150	160	170	180
m120.pep	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETE VVKYRVRRGD					
g120	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETE VVKYRVRRGD					
	130	140	150	160	170	180
	190	200	210	220		
m120.pep	DAVMYFFAPSLNNIPAIQIGYTDDGKTYTLKLKSVQINGQA AKP					
	:					
g120	DTVTYFFAPSLNNIPAIQIGYTDDGKTYTLKLKSVQINGQA AKPX					
	190	200	210	220		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 433>:

a120.seq

```

1  ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGCACT
101 ATTCCGCAG CTACGGCATT CCCGCCACGA TGACATTGTA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAGTCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
251 ATAGAGACAT ACGCAGGGGC AAATGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAACCG AGCAAAGCCC
351 CAAGGCTATG GATTTGTTC ACGCTTGCCTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGTACAGGA AAATACAGCA TAGGCGGCGT
501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
551 TGTATTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACGACG GCAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651 CGGCCAGGCA GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 434; ORF 120.a>:

a120.pep

```

1  MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLHYSGSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAIQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*

```

m120/a120 99.6% identity in 223 aa overlap

	10	20	30	40	50	60
m120.pep	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLXYSYGI PATMTFERSGNAYKIVSTIK					
a120	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLHYSGSYGI PATMTFERSGNAYKIVSTIK					

343

	10	20	30	40	50	60
m120.pep	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
a120	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
m120.pep	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVEVKYRVRRGD					
a120	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVEVKYRVRRGD					
m120.pep	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					
a120	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 435>:

g121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCTGACC GGTTCGCCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATAACAG TTGCCGATTT
351 GCCGCTGCTG GCGGAACTGa cgcggatttt TACCGTCggc gacttcCGCA
401 GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCGCGCA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCCTTCG
551 GCTTCGACAC AGGCGCCGGC AATATGCTGA TGGACgcgtg gacgcaggca
601 cacTGGcagc TGCTTACGA CAAAaaccgt gcAAAGgcgg cacAAGGCAA
651 catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
701 AACCCcacc aaAAAGCACG GGgcGCGaac TgtttgccT AAattggctc
751 gaaacctAcc ttgacggcgg cgaaccga tacgacgtat tgcggacgct
801 ttcccgattc accgcgcaa ccgTttggga cgcgtctca CACGCACGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATtg
1001 cgtggttgC GCGGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMGDGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRKLL
51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 437>:

m121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CCAATTGCTG

```

```

151 GATTTCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTG
201 GCAAGAACTC AGCCGCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCCTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
401 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
451 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
501 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
551 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
601 xxxxxxCAGC TTCCTTACGA CAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCGCAC
701 AACGCCACCC TAAAGCACG GGGCGCGAAC TGTTTGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTTT ACCGCGCAAA CCGTTTGCAG CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTTG
1001 CGTGGTTGGC GCGGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 438; ORF 121>:

m121.pep

```

1 METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRQLL
51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL Axxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
151 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LMDLAECFG TRVSLHSTAD LNLDPPQWEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* menA with menB

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTYPYGRLLRQLLDLQDTGADEL					
g121	METQLYIGIMSGTSMDGADAVLRMDGGKWLGAEGHAFTYPYDRLRRKLLDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEGYSIQLADLPLL					
g121	HRSRMLSQELSRLYAQTAAELLCSQNLAPCDITALGCHGQTVRHAPHEGYSIQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121.pep	AXXX					
g121	AELTRIFTVGDFRSRDLAAGQGAPLVPFHEALFRDDRETRVVLNIGGIANISVLPPGA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
g121	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDAYSHAAADARQMYICDGGIRNPV					
g121	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDAYSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300

345

	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPOQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
	:					
g121	LMADLAECFGTRVSLHSTAE LNLDPOQWVEAAAFWLAACWINRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYY					
g121	GAGYYYY					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 439>:

```

a121.seq
1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGC GGCCGCGC GGACAAGGCG CGCCGCTCGT CCCC GCCTTT
451 CACGAAGCCC TGT TCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCC GACGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTGCGAC
701 AACCACCCC TAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGCTCTA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGDGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWMQA
201 HWQLPYDKNG AKAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

m121/a121 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	:					
a121	METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEHGYISQLADLPLL					
	:					
a121	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPESYSVQLADLPLL					
	70	80	90	100	110	120

346

	130	140	150	160	170	180
m121.pep	XX					
a121	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPFDA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
a121	PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDVASHAAADARQMYICDGGIRNPV					
a121	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFVASHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121.pep	LMADLAECFGRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGRVSLHSTAE LNLDPOWVEAAAFAWMAACWVNRIIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 441>:

m121-1.seq

```

1  ATGGAACAC  AGCTTTACAT  CGGCATCATG  TCGGGAACCA  GCATGGACGG
51  GCGGATGCC  GTACTGATAC  GGATGGACGG  CGGCAAATGG  CTGGGCGCGG
101 AAGGGCACGC  CTTTACCCCC  TACCCCGGCA  GGTACGCCG  CCAATTGCTG
151 GATTTGCAGG  ACACAGGCGC  AGACGAACGT  CACCGCAGCA  GGATTTTGTC
201 GCAAGAACTC  AGCCGCCTAT  ATGCGCAAA  CGCCGCCGAA  CTGCTGTGCA
251 GTCAAAACCT  CGCACCGTCC  GACATTACCG  CCCTCGGCTG  CCACGGGCAA
301 ACCGTCCGAC  ACGCGCCGGA  ACACGGTTAC  AGCATACAGC  TTGCCGATTT
351 GCCGCTGCTG  GCGGAACGGA  CGCGGATTTT  TACCGTCGGC  GACTTCCGCA
401 GCCGCGACCT  TCGGCGCGGC  GGACAAGGCG  CGCCACTCGT  CCCGCCTTT
451 CACGAAGCCC  TGTTCGCGCA  CAACAGGGAA  ACACGCGCGG  TACTGAACAT
501 CGGCGGGATT  GCCAACATCA  GCGTACTCCC  CCCGACGCA  CCGGCCTTCG
551 GCTTCGACAC  AGGGCCGGGC  AATATGCTGA  TGGACGCGTG  GACGCAGGCA
601 CACTGGCAGC  TTCCTTACGA  CAAAAACGGT  GCAAAGGCGG  CACAAGGCAA
651 CATATTGCCG  CAACTGCTCG  ACAGGCTGCT  CGCCACCCG  TATTTGCGAC
701 AACCCACCC  TAAAAGCACG  GGGCGCGAAC  TGTTCGCCCT  AAATTGGGTC
751 GAAACCTACC  TTGACGGCGG  CGAAAACCGA  TACGACGTAT  TCGGACGCT
801 TTCCCGTTTT  ACCGCGCAAA  CCGTTTGCAG  CGCCGTCTCA  CACGCAGCGG
851 CAGATGCCCG  TCAAATGTAC  ATTTGCGGCG  GCGGCATCCG  CAATCCTGTT
901 TTAATGGCGG  ATTTGGCAGA  ATGTTTCGGC  ACACGCGTTT  CCCTGCACAG
951 CACCGCCGAC  CTGAACCTCG  ATCCGCAATG  GGTGGAAGCC  GCCGNATTG
1001 CGTGGTTGGC  GCGGTGTTGG  ATTAATCGCA  TTCCCGGTAG  TCCGCACAAA
1051 GCAACCGGCG  CATCAAACC  GTGTATTCTG  ANCGCGGGAT  ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 442; ORF 121-1>:

m121-1.pep

```

1  METQLYIGIM  SGTSMGDADA  VLIRMDGGKW  LGAEGHAFTP  YPGRLRRQLL
51  DLQDTGADEL  HRSRILSQEL  SRLYAQTAAE  LLCSQNLAPS  DITALGCHGQ
101 TVRHAPHEGY  SIQLADLPL  AERTRIFTVG  DFRSRDLAAG  GQGAPLVPAF
151 HEALFRDNRE  TRAVLNIGGI  ANISVLPFDA  PAFGFDTGPG  NMLMDAWTQA
201 HWQLPYDKNG  AKAAQGNILP  QLLDRLLAHP  YFAQPHPKST  GRELFALNWL
251 ETYLDGGENR  YDVLRTLSRF  TAQTVCDVAV  HAAADARQMY  ICGGGIRNPV
301 LMADLAECFG  TRVSLHSTAD  LNLDPQWVEA  AXFAWLAACW  INRIPGSPHK
351 ATGASKPCIL  XAGYYY*

```

m121-1/g121 95.6% identity in 366 aa overlap

10 20 30 40 50 60

347

```

m121-1.pep  METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTYPGRLLRRQLLDLQDTGADEL
g121         METQLYIGIMSGTSMGDGADAVLVRMDGGKWLGAEGHAFTYPDRLLRRKLLDLQDTGTDEL
              10      20      30      40      50      60

              70      80      90      100     110     120
m121-1.pep  HRSRILSQELSRLYAQTAAELLCSQLNAPSDITALGCHGQTVRHAPHEHGYSIQLADLPLL
g121         HRSRILSQELSRLYAQTAAELLCSQLNAPCDITALGCHGQTVRHAPHEHGYSIQLADLPLL
              70      80      90      100     110     120

              130     140     150     160     170     180
m121-1.pep  AERTRIFTVGDFRSRDLAAGQGQAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
g121         AELTRIFTVGDFRSRDLAAGQGQAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
              130     140     150     160     170     180

              190     200     210     220     230     240
m121-1.pep  PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQHPKST
g121         PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQHPKST
              190     200     210     220     230     240

              250     260     270     280     290     300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDASHAAADARQMYICGGGIRNPV
g121         GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDASHAAADARQMYICGGGIRNPV
              250     260     270     280     290     300

              310     320     330     340     350     360
m121-1.pep  LMADLAECFGTRVSLHSTADNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
g121         LMADLAECFGTRVSLHSTAEINLDPQWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
              310     320     330     340     350     360

m121-1.pep  XAGYYYYX
g121         GAGYYYYX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 443>:

```

a121-1.seq
1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCC TACCCGCGCA GGTTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCGCGGACCT TCGCGCCGCG GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CCGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCGG TATTTCGCAC
701 AACCCACCCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCATTTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAAGGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>:

```

a121-1.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEHGAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQLNAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA

```

201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
 251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
 301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
 351 ATGASKPCIL GAGYYY*

m121-1/a121-1 96.4% identity in 366 aa overlap

m121-1.pep	10	20	30	40	50	60
	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
a121-1	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL					
	10	20	30	40	50	60
m121-1.pep	70	80	90	100	110	120
	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL					
a121-1	HRSRMLSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL					
	70	80	90	100	110	120
m121-1.pep	130	140	150	160	170	180
	AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPDDA					
a121-1	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPDDA					
	130	140	150	160	170	180
m121-1.pep	190	200	210	220	230	240
	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121-1.pep	250	260	270	280	290	300
	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
m121-1.pep	310	320	330	340	350	360
	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAE LNLDPQWVEAAAFAWMAACWVNRI PGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
a121	GAGYYYYX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 445>:

g122.seq

```

1  ATGGCTTTAC TGAGCATCCG CAAGCTGCAC AAACAATACG GCAGCGTAAC
51  CGCCATCCAA TCCTTAGACT TGGACTTGGA AAAAGGCGAA GtcatCGTAC
101 TGCTGGGCCC gTccggtgc ggCAAATCCA CCCTcctgcg ctgcgtaaC
151 GGTTTGGAGC CGCACCAagg cgGCAGCATC GTGATGGACG GTgtcgGCGA
201 ATTCggcAAA GACGTTTCCT GGCAAACCGC CCGGCAAAaA gtcggtatgg
251 tctttcaaag taacgAactg Tttgcccaca tgaccgtcat cgAaaacatc
301 ttcttAggcC CGGTAAagga aCAAAACgc gaccgtgccg aagcaGAGGC
351 gCAAGCCGGC AAactGttgg aacgcgTCGG actgctAGAC CGCAAAAACG
401 CCTATCCGC CGAACTTTCG GCGGTCAGA AACAGCGCAT CGCCATTGTC
451 CGCGCCCTGT GCCTGAATCC GGAAGTCATC CTGCTGGACG AAATCACCGC
501 CGCACTTGAC CCCGAAATGG TGCGGAAGT CTTGGAAGTG GTTTTGAAC
551 TCGCCCGCA AGGGATGAGT ATGCTCATCG TAACCCACGA AATGGGGTTC
601 GCACGCAAAG TTGCCGACCG CATCGTCTTT ATGGACAAAG GCGGCATCGT
651 CCAATCGTCC GACCCGAAA CCTTTTTTTC CGACCAAAA AGCGAACGCG
701 CCCGCCAATT TCTGGCAGGT ATGGACTACT GA
```

This corresponds to the amino acid sequence <SEQ ID 446; ORF 122.ng>:

g122.pep
 1 MALLSIRKLH KQYGSVTAIQ SLDLDLEKGE VIVLLGPSGC GKSTLLRCVN
 51 GLEPHQGGSI VMDGVGEFGK DVSQWTARQK VGMVFQSNEL FAHMTVIENI
 101 FLGPVKEQNR DRAEAEAQAG KLLERVGLLD RKNAYPRELS GGQKQRIAI
 151 RALCLNPEVI LLDEITAALD PEMVREVLEV VLELAREGMS MLIVTHEMGF
 201 ARKVADRIVF MDKGGIVESS DPETFFSAPK SERARQFLAG MDY*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 447>:

m122.seq
 1 GTTGTTCATGA TTAAATCCG CAATATCCAT AAGACCTTTG GCGAAAACAC
 51 TATTTTGC GC GGCATCGATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
 101 TCCTCGGGcC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
 151 GCGTTGGAAG TGCCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
 201 GCTGAAAATC GATTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC
 251 TGCGCCGCAA ATCAKGCATG GTGTTTCAAC AATACAAyCT CTTTCCGCAC
 301 AAAACCGCCT TGGAAAACGT AATGGAAGGA CCGGTTGCCG TACAgGGCAA
 351 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
 401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
 451 CAGCAGCGCG TCGGCATTGC CCGCGCATTG GCGATTCAGC CTGAACTGAT
 501 GCTGTTTAC GAACCGACTT CCGCGCTCGA TCCTGAATTG GTGCAAGATG
 551 TTTTGGATmC CATGAAGGAA TTGGCGCAAG AAGGCTGGAC CATGGTTGTC
 601 GTTACGCATG AAATCAAGTT CGCCTTAGAA GTGGCAACCA CCGwCGTCGT
 651 GATGGACrGC GGCGTTATTG TCGAACCAAG CAGCCCGCAA GATTTGTTTCG
 701 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTTAAGCCA AATCCAATCT
 751 ACCAAGATTT GA

This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:

m122.pep
 1 VVMIKIRNIH KTFGENTILR GIDLVDCKGQ VVILGPSGS GKTTFLRCLN
 51 ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSXM VFQYQNLFPH
 101 KTALENVMEG PVAVQKPAQ QAREEALKLL EKVGLGDKVD LYPYQLSGGQ
 151 QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLDXMKELAQEGWTVV
 201 VTHEIKFALE VATTXVMDX GVIVEQGSPQ DLFDPKHER TRRFLSQIQS
 251 TKI*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng) from *N. gonorrhoeae*:

m122/g122

	10	20	30	40	50	60
m122.pep	VVMIKIRNIHKT	FGENTILRGID	LDVCKGQVVV	ILGPSGSGKT	TFRLCLNALE	MPEDGQI
g122	MALLSIRKLH	KQYGSVTAIQ	SLDLDLEKGE	VIVLLGPSGC	GKSTLLRCVN	GLEPHQGGSI
	10	20	30	40	50	60
	70	80	90	100	110	120
m122.pep	EFDNERPLKID	FSKKPSKHDIL	ALRRKSXMVF	QYQNLFPHK	TALENVMEGP	VAVQKPAQ
g122	VMDGVGEFGK	DVSQWTA-----	RQKVMVFQSN	ELFAHMTVI	ENIFLGPVKE	QNRDRA
	70	80	90	100	110	
	130	140	150	160	170	180
m122.pep	QAREEALKLLE	KVGLGDKVDL	YPYQLSGGQ	QQRVGIARAL	AIQPELMLFD	EPTSALDPEL
g122	EAEAQAGKLL	ERVGLLDRAK	NAYPRELSGG	QKQRIAIVR	ALCLNPEVIL	LDDEITAALD
	120	130	140	150	160	170
	190	200	210	220	230	240
m122.pep	VQDVLDXMKEL	AQEGWTVV	VTHEIKFALE	VATTXVMDX	GVIVEQGSPQ	DLFDPKHER
g122	VREVLEVLEL	AREGMSMLIV	THEMGFARK	VADRIVFMD	KGGIVESSD	PETFFSAPK

350

	180	190	200	210	220	230
	250					
m122.pep	TRRFLSQIQSTKIX					
	: : :					
g122	ARQFLAGMDYX					
	240					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 449>:

```

a122.seq
1   GTTGTCATGA TTAAAATCCG CAATATCCAT AAGACCTTCG GCAAAAATAC
51  CATTTTGC GC GGCATCAATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
101 TCCTCGGGCC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
151 GCGTTGGAAA TGCCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
201 GCTGAAAATC GATTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC
251 TGCGCCGCAA ATCAGGCATG GTGTTTCAAC AATACAACCT CTTTCCGCAC
301 AAAACCGCCT TGGAAAACGT GATGGAAGGA CCGGTTGCCG TACAGGGCAA
351 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
451 CAGCAGCGCG TCGGCATTGC CCGAGCATTG GCGATTGAGC CCGAGCTGAT
501 GTTGTTTGAC GAACCCACTT CCGCGCTTGA CCCCAGATTG GTGCAAGACG
551 TGTGAACGC CATGAAGGAA TTGGCGCGGG AAGGTTGGAC GATGGTCGTC
601 GTTACCCACG AAATCAAGT CGCGCTGGAA GTTGCCACGA CCGTTGTCGT
651 GATGGACGGC GGCGTTATCG TAGAGCAGGG CAGCCCGAAA GAGTTGTTTCG
701 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTAAGCCA AATCCAATCT
751 ACCAAGATTT GA

```

This corresponds to the amino acid sequence <SEQ ID 450; ORF 122.a>:

```

a122.pep
1   VVMIKIRNIH KTFGKNTILR GINLDVCKGQ VVVLGPSPGS GKTTFLRCLN
51  ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSGM VFQQYNLFPH
101 KTALENVM EG PVAVQGKPAA QAREEALKLL EKVGLGDKVD LYPYQLSGGQ
151 QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLNAMKE LAREGWTMVV
201 VTHEIKFALE VATTVVVMDG GVIVEQGS PK ELFDHPKHER TRRFLSQIQS
251 TKI*

```

m122/a122 96.0% identity in 253 aa overlap

	10	20	30	40	50	60
m122.pep	VVMIKIRNIHKTFGENTILRGIDLDVCKGQVVVLGPSPGSGKTTFLRCLNALEMPEDGQI					
a122	VVMIKIRNIHKTFGKNTILRGINLDVCKGQVVVLGPSPGSGKTTFLRCLNALEMPEDGQI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m122.pep	EFDNERPLKIDFSKKPSKHDILALRRKSXMFVQQYNLFPHKTALENVM EGPVAVQGKPAA					
a122	EFDNERPLKIDFSKKPSKHDILALRRKSGMVVQQYNLFPHKTALENVM EGPVAVQGKPAA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m122.pep	QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL					
a122	QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m122.pep	VQDVLDXMKELAQEGWTMVVVTHEIKFALEVATTXVVM DXGVIVEQGSPQDLFDHPKHER					
a122	VQDVLNAMKELAREGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPKELFDHPKHER					
	190	200	210	220	230	240
	250					
m122.pep	TRRFLSQIQSTKIX					

351

a122 |||||
 TRRFLSQIQSTKIX
 250

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 451>:

g122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACCATTTT
51  GCGCGGCATC GATTGATGATG TGGGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCCTCCGG CTGGGTAAAA ACAACATTTT TGGCTGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGCGC GGCCGTTACG
201 CATTGATTTT TCCAAAAAAA CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAGTCCGG AATGGTATTC CAACAATACA ACCTCTTCCC GCATAAAACC
301 GTGTTGGAAA ACGTGATGGA AGGGCCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTTTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGATAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGGTGTCGTA TCGCCCGCGC ACTGGCGATT CAGCCTGAAT TGATGCTGTT
501 TGACGAACCC ACTTCCGCGC TGGACCCGGA GTTGGTGCAA GACGTGTGG
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCACGCT GGAAGTTGCC ACGAACGTCG TCGTGATGGA
651 CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 TCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTGCCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>:

g122-1.pep

```

1  MIKIRNIHKT FGENTILRGI DLDVKGQVV VILGPSGSGK TTFLRCLNAL
51  EMPEDGQIEF DNRPLRIDF SKKTSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMFDEP TSALDPELVQ DVLDMKELA REGWTMVVVT
201 HEIKFTLEVA TNVVVMDGGV IVEQGSPEL FDHLKHERTR RFLSQIQSAK
251 I*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 453>:

m122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA AACTATTTT
51  GCGCGGCATC GATTGATGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCCTCCGG CTGAGGCAAA ACGACGTTT TGGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAATCAGG CATGGTGTTC CAACAATACA ACCTCTTCCC GCACAAAACC
301 GCCTTGGAAA ACGTAATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGCAAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGCGC ATTGGCGATT CAGCCTGAAC TGATGCTGTT
501 TGACGAACCG ACTTCCGCGC TCGATCCTGA ATTGGTGCAA GATGTTTGG
551 ATACCATGAA GGAATTGGCG CAAGAAGGCT GGACCATGGT TGTCGTTACG
601 CATGAAATCA AGTTCGCCTT AGAAGTGGCA ACCACGTCG TCGTGATGGA
651 CGGCGGCGTT ATTGTCGAAC AAGGCAGCCC GCAAGATTG TTCGACCACC
701 CCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>:

m122-1.pep

```

1  MIKIRNIHKT FGENTILRGI DLDVCKGQVV VILGPSGSGK TTFLRCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMFDEP TSALDPELVQ DVLDTMKELA QEGWTMVVVT
201 HEIKFALEVA TTVVVMDGGV IVEQGSQDL FDHPKHERTR RFLSQIQSTK
251 I*

```

m122-1/g122-1 94.8% identity in 251 aa overlap

	10	20	30	40	50	60
m122-1.pep	MIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
g122-1	MIKIRNIHKTFGENTILRGIDLDVKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					

352

	10	20	30	40	50	60
	70	80	90	100	110	120
m122-1.pep	DNERPLKIDFSKKPSKHDILALRRKSGMVFFQYNNLFPHKTALENVMEGPVAVQGKPAAQA					
g122-1	DNERPLRIDFSKKTSKHDILALRRKSGMVFFQYNNLFPHKTVLENVMEGPVAVQGKPAAQA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m122-1.pep	REEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMFLFDEPTSAIDPELVQ					
g122-1	REEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMFLFDEPTSAIDPELVQ					
	130	140	150	160	170	180
	190	200	210	220	230	240
m122-1.pep	DVLDTMKELAQEGWMTMVVTHEIKFALEVATTVVVMDGGVIVEQGSPPQDLFDHPKHHERTR					
g122-1	DVLDTMKELAQEGWMTMVVTHEIKFTLEVATNVVMDGGVIVEQGSPPQDLFDHLKHHERTR					
	190	200	210	220	230	240
	250					
m122-1.pep	RFLSQIQSTKIX					
g122-1	RFLSQIQSAKIX					
	250					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 455>:

a122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTCGGCAAAA ATACCATTTT
51  GCGCGGCATC AATTTGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAATCAGG CATGGTGTTC CAACAATACA ACCTCTTTCC GCACAAAACC
301 GCCTTGGAAG ACGTGATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAGAA GTCGGCTTGG
401 GCGACAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGAGC ATTGGCGATT CAGCCCGAGC TGATGTTGTT
501 TGACGAACCC ACTTCCGCGC TTGACCCCGA GTTGGTGCAA GACGTGTTGA
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCCGCGT GGAAGTTGCC ACGACCGTTG TCGTGATGGA
651 CGCGGCGGTT ATCGTAGAGC AGGCGAGCCC GAAAGAGTTG TTCGACCAAC
701 CCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 456; ORF 122-1.a>:

a122-1.pep

```

1  MIKIRNIHKT FGKNTILRGI NLDVCKGQVV VILGPSGSGK TFLRCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMFLFDEP TSALDPELVQ DVLNAMKELA REGWMTMVVT
201 HEIKFALEVA TTVVMDGGV IVEQGSPEL FDHPKHHERTR RFLSQIQSTK
251 I*

```

a122-1/m122-1 97.2% identity in 251 aa overlap

	10	20	30	40	50	60
a122-1.pep	MIKIRNIHKTFGKNTILRGINLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
m122-1	MIKIRNIHKTFGKNTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
	10	20	30	40	50	60
	70	80	90	100	110	120
a122-1.pep	DNERPLKIDFSKKPSKHDILALRRKSGMVFFQYNNLFPHKTALENVMEGPVAVQGKPAAQA					
m122-1	DNERPLKIDFSKKPSKHDILALRRKSGMVFFQYNNLFPHKTALENVMEGPVAVQGKPAAQA					
	70	80	90	100	110	120
	130	140	150	160	170	180
a122-1.pep	REEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMFLFDEPTSAIDPELVQ					
m122-1	REEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMFLFDEPTSAIDPELVQ					

```
g125.seq
```

seq	1	ATGTCGGGCA	ATGCCTCCTC	TCCTTCATCT	TCCGCCGCCA	TCGGGCTGGT
51	TTGGTTTCGGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACGCTGC	
101	TCGCCCCCTT	GGGCTGGCAG	CGCGGTCTGG	CGGCCCTGCT	TTTGGGTCAT	
151	GCCGTCGGCG	GCGCGCTGTT	TTTTGCGGCG	GCGTATATCG	GCGCACTGAC	
201	CGGACGCAGC	TCGATGGAAA	GTGTGCGCCT	GTCGTTCCGG	AAATGCGGTT	
251	CAGTGCTGTT	TTCCGTGGCG	AATATGCTGC	AAC TGCCCGG	CTGGACGGCG	
301	GTGATGATTT	ACGTCGGCGC	AacggTCAGC	TCCGCTTTGG	GCAAAGTGTT	
351	GTGGGACggc	gaATCCTTTG	TCTGGTGGGC	ATTGGCAAAC	GGCGCACTGA	
401	TCGTGCTGTG	GCTGGTTTTC	GGCGCACGCA	GAACGGGCGG	GCTGAAAACC	
451	GTTTCGATGC	TGCTGATGCT	GCTTGCCGTG	TTGTGGTTGA	GCGTCGAAGT	
501	GTTTCGTTTCG	TCCGGCACAA	ACGCCGCGCC	CGCCGTTTTCA	GACGGCATGA	
551	CCTTCGGAAC	GGCAGTCGAA	CTGTCCGCCG	TCATGCCGCT	TTCTGGCTG	
601	CCGCTGGCCG	CCGATACAC	CGGCCAAGCA	CGCCGCCCGT	TTGCGGCAAC	
651	CCTGACGGCA	ACGCTCGCCT	ATACGCTGAC	GGGCTGCTGG	ATGTATGCCT	
701	TGGGTTTGGC	GGCGGCTCTG	TTTACCGGAG	AAACCGACGT	GGCGAAAATC	
751	CTGTTGGGCG	CGGGCTTGGG	CATAACGGGC	ATTCTGGCAG	TCGTCTCTC	
801	CACCGTTACC	ACAACGTTTC	TCGATACCTA	TTCCGCCGGC	GCGAGTGCGA	
851	ACAACATTTC	CGCGCGTTTT	GCGGAAATAC	CCGTCGCTGT	CGGCGTTACC	
901	CTGAtccgca	cggtgcttgc	cgtcatgtc	cccgttaccg	aatataaaaa	
951	cttctctgctg	cttatccgct	cggtatgttg	gccgatggcg	ggtggttttg	
1001	atttgcgaCT	TTTTttgctt	AAAACGGCGT	GA		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 459>:

m125.seq

seq	ATGTCGGGCA	ATGCCTCCTC	TCCTTCATCT	TCCTCCGCCA	TCGGGCTGAT
51	TTGGTTCGGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACGCTGC
101	TTGCGCTTT	GGGCTGGCAG	CGCGTCTGG	CGGCTCTACT	TTTGGGT CAT
151	GCCGTCGGG	GCGCGCTGT	TTTTGCGGCG	GCGTATATCG	GCGCACTGAC
201	CGGACGCAGC	TCGATGAAA	CGCTGCGCT	TCGTTTCGGC	AAACGCGTT
251	CAGTGTGTT	TTCCGTGGCG	AATATGCTGC	AACTGGCCGG	CTGGACGGCG
301	GTGATGATT	ACGCCGGCGC	AACGGTCAGC	TCCGCTTTGG	GCAAAGTGTT
351	GTGGGACGGC	GAATCTTTTG	TCTGTTGGGC	ATTGCAAAAC	GGCGCGCTGA
401	TTGTGCTGTG	GCTGGTTTTC	GGCGCACGCA	AAACAGCGCG	GCTGAAAACC
451	GTTTCGATCG	TGCTGATGCT	GTTGGCGGTT	CTGTGGCTGA	GTGCCGAAGT
501	CTTTTCCAGC	GCAAGGACCA	CCGCCGCACA	GGTTTCAGAC	GCTCATGAGTT
551	TCGGAACGCG	AGTCGAGCTG	TCCGCCGTGA	TGCCGCTTTC	CTGGCTGCCG
601	CTTGCCGCGC	ACTACACGCG	CCACGCGCGC	CGCCCGTTTG	CGGCAACCCT
651	GACGGCAACG	CTCGCCTACA	CGCTGACCGG	CTGCTGGATG	TATGCCTTGG
701	GTTTGGCAGC	GGCGTGTTC	ACCGGAGAAA	CCGACGTGGC	AAAAATCTGT
751	CTGGGCGCAr	GTTTGtGTGC	GGCAGGCATT	TGGCGGTGCG	TCCTCTCCAC
801	CGTTACCACA	ACGTTTCTCG	ATGCCTATTG	CGCCGCGCGC	AGTGCGAACA

354

851 ACATTTCCGC GCGTTTTGCG GAAACACCCG TCGCTGTCrG CGTTACCCCTG
 901 ATCGGCACGG TACTTGCCGT CATGCTGCCC GTTACCGAAT ATGAAAACTT
 951 CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCGGgC GGTTTTGATT
 1001 GCCGACTTTT TCGTCTTGAA ACGGCGTGA

This corresponds to the amino acid sequence <SEQ ID 460; ORF 125>:

m125.pep
 1 MSGNASSPSS SSAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
 51 AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSA NMLQLAGWTA
 101 VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
 151 VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
 201 LAADYTRHAR RPFATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL
 251 LGAXLGAAGI LAVVLSTVTT TFLDAYSAGA SANNISARFA ETPVAVXVTL
 301 IGTVLAVMLP VTEYENFLLL IGSVFAPMAG GFDCRLFRLE TA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 125 shows 92.1% identity over a 343 aa overlap with a predicted ORF (ORF 125.ng) from *N. gonorrhoeae*:

m125/g125

m125.pep	MSGNASSPSSSSAIGLIWFGAAVVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
g125	MSGNASSPSSSAAIGLVWFGAAVVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
	10 20 30 40 50 60
m125.pep	AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
g125	AYIGALTGRSSMESVRLSFGKCGSVLFVANMLQLAGWTAVMIYVGATVSSALGKVLWDG
	70 80 90 100 110 120
m125.pep	ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQ-VS
g125	ESFVWWALANGALIVLWLVFGARRTGGLKTVSMLLMLLAVLWLSVEVFASSTNAAPAVS
	130 140 150 160 170 179
m125.pep	DGMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAAL
g125	DGMTFGTAVELSAVMPLSWLPLAADYTRQARRPFAATLTATLAYTLTGCWMYALGLAAAL
	180 190 200 210 220 230 239
m125.pep	FTGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVT
g125	FTGETDVAKILLGAGLGITGILAVVLSTVTTTFLDTYSAGASANNISARFAEIPVAVGVT
	240 250 260 270 280 290 299
m125.pep	LIGTVLAVMLPVTEYENFLLIGSVFAPMAGGFDCRLFRLETAX
g125	LIRTVLAVMLPVTEYKNFLLIRSVFGPMAGGFDCRLFCLKTAX
	300 310 320 330 340

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 461>:

a125.seq
 1 ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGAT
 51 TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACACTGC
 101 TTGCGCCTTT GGGCTGGCAG CGCGGTCTGG CGGCTCTGCT TTTGGGTCAT

a125.pep

1	MSGNASSPSS	SAAIGLIWFG	AAVSIAEIST	GTLLAPLGWQ	RGLAALLLGH
51	AVGGALFFAA	AYIGALTGRS	SMESVRLSFG	KRGSVLFSVA	NMLQLAGWTA
101	VMIYAGATVS	SALGKVLWDG	ESFVWWALAN	GALIVLWLWF	GARKTGGLKT
151	VSMLMLLLAV	LWLSAEVFS	AGSTAAQVSD	GMSFGTAVEL	SAVMPLSVLP
201	LAADYTRHAR	RPFAATLTAT	LAYLTGCGWM	YALGALAAAL	TGETDVAKIL
251	LGAGLGAAGI	LAVVLSTVTT	TFLLDAYSAGV	SANNISAKLS	EIPIAVAVAV
301	VGTL LAVLLP	VTEYENFLLL	IGSVFAPMAX	GFDCLRFRL	TA*

	10	20	30	40	50	60
m125.pep	MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQQRGLAALLLGHAVGGALFFFAA					
a125	:					
	10	20	30	40	50	60
	70	80	90	100	110	120
m125.pep	AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVWLWDG					
a125						
	70	80	90	100	110	120
	130	140	150	160	170	180
m125.pep	ESFVWWALANGALIVLWLVFGARKTGGLKTVMMLMLLAVLWLSAEVFSTAGSTAQQVSD					
a125						
	130	140	150	160	170	180
	190	200	210	220	230	240
m125.pep	GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAALF					
a125						
	190	200	210	220	230	240
	250	260	270	280	290	300
m125.pep	TGETDVAKILLGAXLGAAGILAVVLSTVTTFLDAYSAGASANNISARFAETPVAVXVTL					
a125						
	250	260	270	280	290	300
	310	320	330	340		
m125.pep	IGTVLAVMLPVTEYENFLLIGSVFAPMAGGFDCRLFRLLETAX					
a125	: :					
	VGTLLAVLLPVTEYENFLLIGSVFAPMAXGFDCRLFRLLETAX					

310 320 330 340

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 463>:

g126.seq

```

1   AtgccgctcTG AAaccCcaaa ggcACGCCGC CGGCTTTCAG ACGGCATCGC
51  GTCCGACAAC CATACCAAAG AATCCATCAT GCTCACCCtg tacggcGAAA
101 CTTTCCCTTC GCGGCTGCTg ctcggcacgG eggcctacCC GACCCCTGAA
151 ATCCTCAAAC AATCCGTCCG AACGCCCGG CCCGCGATGA ttaccGTCTC
201 GCTGCGCCGC ACGGGATGCG GCGGCGAGGC GCACGGTCAG GGGTTTTGGT
251 CGTGCTTCA AGAAACCGGC GTTCCCGTCC TGCCGAACAC GGCAGGCTGC
301 CAAAGCGTGC AGGAAGCGGT AACGACGGCG CAAATGGCGC GCGAAGTGT
351 TGAAACCGAT TGGATAAAAT TGGAACTCAT CGGCGACGAC GACACCTTGC
401 AGCCGGACGT GTTCCAACTC GTCGAAGCGG CGGAAATCCT GATTAAAGAC
451 GGCTTCAAAG TGCTGCCTTA TTGCACCGAA GACCTGATTG CCTGCCGCCG
501 CCTGCTCGAT GCGGGCTGTC AGGCGTTGAT GCCGTGGGCG GCTCCCATCG
551 GCACGGGTTT GGGGCGGGT CACGCCTATG CGTCAAAAT CCTGCGCGAA
601 CGCTGCCCCG ACACGCCGCT GATTATCGAC GCGGGCTTGG GTTTGCCCTC
651 CCAAGCGGCA CAAGTGATGG AATGGGGTTT TGACGGCGTA TTGTTAAACA
701 CCGCCGTTTC CCGCAGCGGC GACCCCGTCA ACATGGCGCG CGCCTTCGCA
751 CTCGCCGTCG AATCCGGACG GCTGGCATTG GAAGCCGGGC CGGTCGAAGC
801 GCGAACCAAA GCCCAAGCCA GCACGCCGAC AGTCGGACAA CCGTTTTGGC
851 ATTCGGCGGA ATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 464; ORF 126.ng>:

g126.pep

```

1   MPSETPKARR RLSDGIASDN HTKESIMLT YGETFPSRL LGTAAYPTPE
51  ILKQSVRTAR PAMITVSLRR TGCGGEAHGQ GFWSLLQETG VPVLPNTAGC
101 QSVQEAVTTA QMAREVFETD WIKLELIGDD DTLQPDVFQL VEA AEILIKD
151 GFKVLPYCTE DLIACRRLD AGCQALMPWA APIGTGLGAV HAYALKILRE
201 RLPDTPLIID AGLGLPSQAA QVMWGFQDV LLNTAVSRSG DPNMARAF
251 LAVESGRLAF EAGPVEARTK AQA STPTVGQ PFWSAEY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 465>:

m126.seq (partial)

```

1   ..CACTATACAA AGGAACCCAT TATGCTCACC CTATACGGCG AAAC TTTCCC
51  CTCGCGGCTG CTGCTCGGCA CGGCTGCCTA CCCGACCCCG GAAATCCTCA
101 AACAAATCCAT CCAAACCGCC CAGCCTGCGA TGATTACCGT CTCGCTGCGC
151 CGCGCGGGAA GCGGCGGCGA GGCGCACGGT CAGGGGTTTT GGTCGCTGCT
201 TCAAGAAACC GCGGTTCCCG TCCTGCCGAA CACGGCAGGC TGCCAAAGCG
251 TGCAGGAAGC GGTAACGACG GCGCAAATGG CGCGCGAAGT GTTTGAAACC
301 GATTGGATAA AATTGGAAC TATCGGAGAT GACGACACCT TGCAGCCGGA
351 TGTGTTCCAG CTTGTCGAAG CGGCGGAAAT CCTGATTAAA GACGGCTTCA
401 AAGTGCTGCC TTATTGCACC GAAGACCTGA TTGCCTGCCG CCGCTGCTC
451 GACGCGGGCT GTCAGGCGTT GATGCCGTGG GCGGCCCCGA TCGGCACGGG
501 TTTGGGCGCG GTTCACGCCT ACGCGTTGAA CGTCCTGCGC GAACGCCTGC
551 CCGACACGCC GCTGATTATC GACGCGGGCT TGGGTTTGCC CTCACAGGCG
601 GCACAACTGA TGGAATGGGG CTTTGACGGC GTGCTTTTGA ATACTGCCGT
651 TTCCCGCAGC GGCGATCCGG TCAATATGGC ACGCGCCTTC GCACTCGCCG
701 TCGAATCCGG ACGGCTGGCA TTTGAAGCCG GACCGGTCGA AGCAGCGCAC
751 AAAGCGCAAG CCAGCACGCC GACAGTCGGA CAACCGTTTT GGCATTCCGG
801 GGAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 466; ORF 126>:

m126.pep (partial)

```

1   ..HYTKPEIMLT LYGETFPSRL LLGTAAYPTP EILKQSIQTA QPAMITVSLR
51  RAGSGGEAHG QGFWSLLQET GVPVLPNTAG CQSVQEAVTT AQMAREVFET

```

```

101 DWIKLELIGD DDTLQPDVFQ LVEAAEILIK DGFKVLPYCT EDLIACRRL
151 DAGCQALMPW AAPIGTGLGA VHAYALNVLR ERLPDTPLII DAGLGLPSQA
201 AQVMEWGFDG VLLNTAVSRS GDPVNMARAF ALAVESGRLA FEAGPVEARD
251 KQASTPTVG QPFWHSAEY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from *N. gonorrhoeae*:

m126/g126

```

                                10      20      30      40
m126.pep                      HYTKPEIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQ
                                ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      MPSETPKARRRLSDGIASDNHTKESIMLTLYGETFPSRLLLGTAAYPTPEILKQSVRTAR
                                10      20      30      40      50      60

                                50      60      70      80      90      100
m126.pep      PAMITVSLRRAGSGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      PAMITVSLRRTGCGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                70      80      90      100      110      120

                                110      120      130      140      150      160
m126.pep      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                130      140      150      160      170      180

                                170      180      190      200      210      220
m126.pep      APIGTGLGAVHAYALNVLRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSG
                ||||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      APIGTGLGAVHAYALKILRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSG
                190      200      210      220      230      240

                                230      240      250      260      270
m126.pep      DPVNMARAFALAVESGRLAFEAGPVEARDKAQASTPTVGQPFWHSAEYX
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      DPVNMARAFALAVESGRLAFEAGPVEARTKAQASTPTVGQPFWHSAEYX
                250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 467>:

```

a126.seq
1   TTGTTAATCC ACTATACAAA GGAACCCATT ATGCTCACCC TGTACAGCGA
51  AACTTTCCCT TCGCGGCTGC TGCTCGGCAC AGCCGCCTAC CCGACCCCTG
101 AAATCCTCAA ACAATCCGTC CGAACCGCCC GGCCCGCGAT GATTACCGTC
151 TCGCTGCGCC GCGCGGGATG CGGCGGCGAG GCGCACGGTC AGGGGTTTTG
201 GTCGCTGCTT CAAGAAACCG GCGTTCCTCGT CCTGCCGAAC ACGGCAGGCT
251 GCCAAAGCGT GCAGGAAGCG GTAACGACGG CGCAAATGGC GCGCGAAGTG
301 TTTGAAACCG ATTGGATTAA ACTCGAATC ATCGGCGACG ACGACACCTT
351 GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC GGCGGAAATC CTGATTAAAG
401 ACGGCTTCAA AGTGCTGCCT TATTGCACCG AAGACCTGAT TGCCTGCCGC
451 CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG ATGCCGTGGG CGGCCCGGAT
501 CGGCACGGGT TTGGGCGCGG TTCACGCCTA CGCGTTGAAC GTCCTGCGCG
551 AACGCCTGCC CGACACGCCG CTGATTATCG ACGCGGGCTT GGGTTTGCCC
601 TCACAGGCGG CACAAGTGAT GGAATGGGGC TTGACGGCG TGCTTTTGAA
651 TACTGCCGTT TCCCGCAGCG GCGATCCGGT CAATATGGCA CGCGCCTTCG
701 CACTCGCCGT CGAATCCGGA CGGCTGGCAT TTGAAGCCGG ACCGGTCGAA
751 GCACGCGACA AAGCGCAAGC CAGCACGCCG ACAGTCGGAC AACCGTTTTG
801 GCATTGCGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 468; ORF 126.a>:

```
a126.pep
  1  LLHYTKPEI MLTYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV
 51  SLRRAGCGGE AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV
101  FETDWIKLEL IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR
151  RLLDAGCQAL MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP
201  SQAAQVMWEG FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE
251  ARDKAQASTP TVGQPFWHS EY*
```

m126/a126 98.1% identity in 269 aa overlap

```

              10      20      30      40      50
m126.pep      HYTKPEIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGE
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           LLHYTKPEIMLTLYSETFPSRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRAGCGGE
              10      20      30      40      50      60

              60      70      80      90      100     110
m126.pep      AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREV FETDWIKLELIGDDDTLQPD
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREV FETDWIKLELIGDDDTLQPD
              70      80      90      100     110     120

              120     130     140     150     160     170
m126.pep      VFQVLEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           VFQVLEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN
              130     140     150     160     170     180

              180     190     200     210     220     230
m126.pep      VLRERLPDTP LIIDAGLGLPSQAAQVMWEGFDGVLLNTAVSRSGDPVNMA RAFALAVESG
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           VLRERLPDTP LIIDAGLGLPSQAAQVMWEGFDGVLLNTAVSRSGDPVNMA RAFALAVESG
              190     200     210     220     230     240

              240     250     260     270
m126.pep      RLAFEAGPVEARDKAQASTPTVGQPFWHS EYX
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           RLAFEAGPVEARDKAQASTPTVGQPFWHS EYX
              250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 469>:

```
g126-1.seq
  1  ATGCTCAGCC TGTACGGCGA AACTTTCCT TCGCGGCTGC TGCTCGGCAC
 51  GGCCGCCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACC GCCC
101  GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCACGGGATG CGGCGGCGAG
151  GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201  CCTGCGGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251  CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAATC
301  ATCGGCGACG ACGACACCTT GCAGCCGGAC GTGTTCCAAC TCGTCGAAGC
351  GCGGGAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401  AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ATGCGGGCTG TCAGGCGTTG
451  ATGCCGTGGG CGGCTCCCAT CGGCACGGGT TTGGGGGCGG TTCACGCCTA
501  TGCGCTCAAA ATCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
551  ACGCGGCTT GGGTTTGCTT TCCCAAGCGG CACAAGTGAT GGAATGGGGT
601  TTTGACGGCG TATTGTAAAC CACCGCCGTT TCCCGCAGCG GCGACCCCGT
651  CAACATGGCG CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701  TTGAAGCCGG GCCGGTCGAA GCGCGAACCA AAGCCCAAGC CAGCACGCCG
751  ACAGTCGGAC AACCGTTTTG GCATTGCGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 470; ORF 126-1.ng>:

```
g126-1.pep
  1  MLTYGETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRTGCGGE
```

```

51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251 TVGQPFWHSA EY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 471>:

m126-1.seq

```

1 ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
51 GGCTGCCTAC CCGACCCCGG AAATCCTCAA ACAATCCATC CAAACCGCCC
101 AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CCGCGGCGAG
151 GCGCACGGTC AGGGGTTTGT GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAACCTC
301 ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTTCCAGC TTGTCGAAGC
351 GCGCGAAATC CTGATTAAAG ACGGCTTCAA AGTGTGCTT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CCGGTTGAAC GTCCTGCGCG AACGCCGCGC CGACACGCGC CTGATTATCG
551 ACGCGGGCTT GGGTTTGGCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751 ACAGTCGGAC AACCGTTTGT GCATTGCGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>:

m126-1.pep

```

1 MLTYGETFP SRLLLGTAA YPTPEILKQSI QTAQPAMITV SLRRAGSGGE
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWHSA EY*

```

m126-1/g126-1 96.9% identity in 262 aa overlap

	10	20	30	40	50	60
m126-1.pep	MLTYGETFP	SRLLLGTAA	YPTPEILKQSI	QTAQPAMITV	SLRRAGSGGE	AHGQGFWSLL
g126-1	MLTYGETFP	SRLLLGTAA	YPTPEILKQSV	RTARPAMITV	SLRRTGCGGE	AHGQGFWSLL
	70	80	90	100	110	120
m126-1.pep	QETGVPVLP	NTAGCQSVQEA	VTTAQMAREV	FETDWIKLEL	IGDDDTLQPD	VFQVLEAAEI
g126-1	QETGVPVLP	NTAGCQSVQEA	VTTAQMAREV	FETDWIKLEL	IGDDDTLQPD	VFQVLEAAEI
	130	140	150	160	170	180
m126-1.pep	LIKDGFKVLP	YCTEDLIACR	RLLDAGCQAL	MPWAAPIGTG	LGAVHAYALN	VLRERLPDTP
g126-1	LIKDGFKVLP	YCTEDLIACR	RLLDAGCQAL	MPWAAPIGTG	LGAVHAYALK	ILRERLPDTP
	190	200	210	220	230	240
m126-1.pep	LIIDAGLGLP	SQAAQVMEWG	FDGVLLNTAV	SRSGDPVNMA	RAFALAVESG	RLAFEAGPVE
g126-1	LIIDAGLGLP	SQAAQVMEWG	FDGVLLNTAV	SRSGDPVNMA	RAFALAVESG	RLAFEAGPVE
	250	260				
m126-1.pep	ARDKAQASTP	TVGQPFWHSAEYX				
g126-1	ARTKAQASTP	TVGQPFWHSAEYX				
	250	260				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 473>:

a126-1.seq

```

1 ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
51 AGCCGCGCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101 GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CCGCGGCGAG

```

```

151 GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATTAA ACTCGAACTC
301 ATCGGCGACG ACGACACCTT GCAGCCGGAT GTGTTCCAAC TTGTGGAAGC
351 GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCGG CTGATTATCG
551 ACGCGGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751 ACAGTCGGAC AACC GTTTTGCATTTCGGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 474; ORF 126-1.a>:

a126-1.pep

```

1  MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRAGCGGE
51  AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLNTAV SRS GPDVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWSHA EY*

```

a126-1/m126-1 98.1% identity in 262 aa overlap

	10	20	30	40	50	60
a126-1.pep	MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRAGCGGE AHGQGFWSLL					
m126-1	MLTLYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE AHGQGFWSLL					
	10	20	30	40	50	60
	70	80	90	100	110	120
a126-1.pep	QETGVPVLP NTAGCQSVQEA VTTAQMAREV FETDWIKLEL IGDDDTLQPD VFQLVEAAEI					
m126-1	QETGVPVLP NTAGCQSVQEA VTTAQMAREV FETDWIKLEL IGDDDTLQPD VFQLVEAAEI					
	70	80	90	100	110	120
	130	140	150	160	170	180
a126-1.pep	LIKDGFKVLP YCTEDLIACR RLLDAGCQAL MPWAAPIGT GLGAVHAYALN VLRERLPDTP					
m126-1	LIKDGFKVLP YCTEDLIACR RLLDAGCQAL MPWAAPIGT GLGAVHAYALN VLRERLPDTP					
	130	140	150	160	170	180
	190	200	210	220	230	240
a126-1.pep	LIIDAGLGLP SQAAQVMEWG FDGVLNTAV SRS GPDVNMA RAFALAVESG RLAFEAGPVE					
m126-1	LIIDAGLGLP SQAAQVMEWG FDGVLNTAV SRS GPDVNMA RAFALAVESG RLAFEAGPVE					
	190	200	210	220	230	240
	250	260				
a126-1.pep	ARDKAQASTP TVGQPFWSHA EYX					
m126-1	ARDKAQASTP TVGQPFWSHA EYX					
	250	260				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 475>:

gl27.seq

```

1  ATGGAATAT GGAATATGTT GAACACTTGG CCCGATGCCG TCCCGATACG
51  CGCGGAGGCG GCCGAATCCG TGGCGGCGGT CGCGGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAGAC GGCATCCGGA TTTCGGCATC
151 GAAAGCAAGC GCGGTTTTTT GGTGCCAGC CGCAATATAA CGTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGTC GCGCGAAATT CAAACGCTGG
251 CTTTGTGCGAT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACAAAAGAA
301 CTGATTATGT GTCTGTGCGG CAGTATTTTA aggtctGCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCC
401 ACATCAATCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA

```

361

```

551  CGGTCGAAAT  CCCCGTTCCC  ATCCATTGG  ATTCGGATGA  AGCCGTATGC
601  CGTCTGAAAG  CCGTACTCGA  GCCCTTGTGC  GCGCCCTACA  TCCCCGCCAT
651  TCAGCGGTAT  TTGGAAAACG  TGCAGGCGGA  AAAACTGTTT  ATCACGCCCG
701  CCGCCAGGCC  GCGCGTTACC  CGCGTACCGT  ACGACGACAA  GGCATACCGC
751  ATCATCGTCC  GCTTCGCCTC  CCCCGTTTCA  AAGCGGCTGG  AAATCCAACA
801  GGCGGTTATG  GACGAATTTT  TGCGCGTACA  ATACCGCCTG  TTAAATCATC
851  CCGCCGgctc  cgAAACACTT  TAA

```

This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:

```

g127.pep
  1  MEIWNMLNTW  PDAVPIRAEA  AESVAAVAAL  LLARALLLNI  HFRRHPDFGI
 51  ESKRRFLVAS  RNITLLLVL  SLAFIWSAQI  QTLALSMFAV  AAHVVVATKE
101  LIMCLSGSIL  RSATQQYSVG  DYIEINGLRG  RVVDINLLNT  LMMQVGPNPL
151  VGQLAGTTVS  FPNSLLLSHP  VRRDNILGDY  VIHTVEIPVP  IHLDSDEAVC
201  RLKAVLEPLC  APYIPAIQRY  LENVQAEKLF  ITPAARPRVT  RVPYDDKAYR
251  IIVRFASPVS  KRLEIQQAVM  DEFLRVQYRL  LNHPAGSETL  *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 477>:

```

m127.seq
  1  ATGGAAATAT  GGAATATGTT  GGACACTTGG  CTCGGTGCCG  TCCCGATACG
 51  TCGCGAGGCG  GTCGAATCCG  TGGCGGCGGT  TCGCGCTTGT  CTGCTGGCGC
101  GCGCCCTTCT  GTTGAATATC  CACTTCAAAC  GGCATCCGGA  TTTCGGCATC
151  GAAAGCAAGC  GGCGGTTTTT  GGTTGCCAGC  CGCAATATAA  CGCTGCTTTT
201  GGTGCTGTTT  TCGCTGGCAT  TTATCTGGTC  GGCGCAAATC  CAAACGCTGG
251  CTTTGTGCGAT  GTTTGCGGTG  GCGGCGGCGG  TCGTCGTGGC  GACGAAGGAA
301  CTGATTATGT  GTCTGTGCGG  CAGTATTTTA  AGGTCTGCCA  CCCAGCAATA
351  CTCGGTCGGC  GACTATATCG  AAATCAACGG  CCTGCGCGGG  CGCGTGGTCG
401  ACATCAACCT  GTTGAACACG  CTGATGATGC  AGGTGCGTCC  GAACCCCTTG
451  GTCGGACAGC  TTGCGGGAAC  CACCGTTTCT  TTCCCCAACA  GCCTGTTGTT
501  GAGCCACCCC  GTGCGCCGCG  ACAATATTTT  GGGCGACTAT  GTCATCCATA
551  CGGTCGAAAT  CCCCGTTCCC  ATCCATTGG  ATTCGGATGA  AGCCGTATGC
601  CGTCTGAAAG  CCGTACTCGA  GCCCTTGTGC  GCGCCCTACA  TCCCCGCCAT
651  CCAACGGsAT  TTGGAAAACG  TGCAGGCGGA  AAAACTGTTT  ATCACGCCCG
701  CCGCCAGACC  GCGCGTTACC  CGCGTGCCGT  ACGATGACAA  GGCATACCGC
751  ATCATCGTCC  GCTTCGCTTC  CCCCGTTTCA  AAGCGGCTGG  AAATCCAACA
801  GGCGGTTATG  GACGAATTTT  TGCGCGTACA  ATACCGCCTG  TTAAATCACC
851  CCGCCGCTC  CGAAACACTT  TAA

```

This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:

```

m127.pep
  1  MEIWNMLDTW  LGAVPIRAEA  VESVAAVAAL  LLARALLLNI  HFKRHPDFGI
 51  ESKRRFLVAS  RNITLLLVL  SLAFIWSAQI  QTLALSMFAV  AAHVVVATKE
101  LIMCLSGSIL  RSATQQYSVG  DYIEINGLRG  RVVDINLLNT  LMMQVGPNPL
151  VGQLAGTTVS  FPNSLLLSHP  VRRDNILGDY  VIHTVEIPVP  IHLDSDEAVC
201  RLKAVLEPLC  APYIPAIQRY  LENVQAEKLF  ITPAARPRVT  RVPYDDKAYR
251  IIVRFASPVS  KRLEIQQAVM  DEFLRVQYRL  LNHPAGSETL  *

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng) from *N. gonorrhoeae*:

m127/g127

```

          10      20      30      40      50      60
m127.pep  MEIWNMLDTWLGAVPIRAEAESVAAVAALLLARALLLNIHFKRHPDFGIESKRRFLVAS
          |||||:|  |||||:|||||||||||||||||||||:|||||||||
g127       MEIWNMLNTWPDVPIRAEAESVAAVAALLLARALLLNIHFRRHPDFGIESKRRFLVAS
          10      20      30      40      50      60

          70      80      90     100     110     120
m127.pep  RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG
          |||||:|  |||||:|||||||||||||||||||||:|||||||||
g127       RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG

```


362

	70	80	90	100	110	120
	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
g127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSDCAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
g127	VIHTVEIPVPIHLDSDCAVCRLKAVLEPLCAPYIPAIQRYLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFRLVQYRLNHPAGSETLX					
g127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFRLVQYRLNHPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 479>:

```

a127.seq
1  ATGGAAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
51  TCGCGAGGCG GTCGAATCCG TGGCGGTGGT CGCGGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTCGGCATC
151 GAAAGCAAGC GCGCGTTTTT GGTGCCAGC CGCAATATAA CGCTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GCGCGCAAATC CAAACGCTGG
251 CTTTGTGCGAT GTTTGCGGTG GCGGCGCGCG TCGTCGTGGC GACGAAGGAA
301 CTGATTATGT GTCTGTCGGG CAGCATTTTA AGGTCTGCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAACCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAC GTCATCCATA
551 CCGTCGAAAT CCCGGTTCCT ATCCATTGGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTC GCGCCCTACA TCCCGGCCAT
651 CCAACGGCAT TTGGAACACG TGCAGGCGGA AAAACTGTTT ATCAGCCCGG
701 CCGCCAAACC GCGCGTTACC CGCGTGCCGT ACGATGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCCTC CCCCCTTTCA AAGCGGCTGG AAATCCAACA
801 GCGCGTTATG GACGAATTTT TGC GCGTACA ATACCGCCTG TTAATTACC
851 CCGCCGCTC CGAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 480; ORF 127.a>:

```

a127.pep
1  MEIWNMLDTW LGAVPIRAEA VESVAVVAAL LLARALLNI HFKRHPDFGI
51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAHVVVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPV IHLSDCAVC
201 RLKAVLEPLC APYIPAIQRH LENVQAEKLF ITPAAKPRVT RVPYDDKAYR
251 IIVRFASPVSKRLEIQQAVM DEFRLVQYRL LNPAGSETL *

```

m127/a127 98.6% identity in 290 aa overlap

	10	20	30	40	50	60
m127.pep	MEIWNMLDTWLGAVPIRAEAVESVAVVAALLLARALLNIHFKRHPDFGIESKRRFLVAS					
a127	MEIWNMLDTWLGAVPIRAEAVESVAVVAALLLARALLNIHFKRHPDFGIESKRRFLVAS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m127.pep	RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG					
a127	RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG					
	70	80	90	100	110	120

	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNILGDY					
a127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSEAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
a127	VIHTVEIPVPIHLDSEAVCRLKAVLEPLCAPYIPAIQRHLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNHPAGSETLX					
a127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNYPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 481>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
51  aatccaaacc gaagACatca AACCCGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACCGCTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
401 TCGCGGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCAG ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCCTC
501 CCAAACCGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAACACCTCG GTCTCGCCGA CCCGAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTTCTGGCAG GCCTGTTCGC
1101 CCAAATCAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCC
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTaca AAGGCCGCCG CCGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCAGCC TCTTCCACGA
1401 AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAAcggcgta GAATGGGACG CGGTGCAACT GCCCAGCCAG
1501 TTTATGGAAG ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgCGGAGC CCCTGCCGAA AGAACTCTC GACAAAATGC
1601 TcgcCGCCAA AAATCTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCct
1851 cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCCGAGC
1951 gcgGCGGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC

```

2001 ACTGCTGCGC CAAagcggT TCGACAACGC gGcttga

This corresponds to the amino acid sequence <SEQ ID 482; ORF 128.ng>:

```
g128.pep
1  MIDNALLHLG  EEPRFNQIQT  EDIKPÄVQTA  IAEARGQIAA  VKAQHTGWGA
51  NTVERLTGIT  ERVGRIWGVV  SHLNSVVDTP  ELRAVYNELM  PEITVFFTEI
101 GQDIELYNRF  KTIKNSPEFA  TLSPAQKTKL  DHDLRDFVLS  GAELPPERQA
151 ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201 AAQSEKGTGY  KIGLQIPHYL  AVIQYAGNRE  LREQIYRAYV  TRASELSNDG
251 KFDNTANIDR  TLENALKTAK  LLGFKNYAEL  SLATKMADTP  EQVLNPLHDL
301 ARRAKPYAEK  DLAEVKAFAR  EHLGLADPQP  WDLSYAGEKL  REAKYAFSET
351 EVKKYFPVGK  VLAGLFAQIK  KLYGIGFAEK  TVPVWHKDVR  YFELQQNGKT
401 IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFA  DGTLLQLPTAY  LVCNFAPPVG
451 GKEARLSHDE  ILTLFHETGH  GLHLLTQVD  ELGVSGINGV  EWDAVELPSQ
501 FMENFVWEYN  VLAQMSAHEE  TGEPLPKELF  DKMLAAKNFQ  RGMFLVRQME
551 ALFDMMIYS  ESDECRLKNW  QQVLDSVRKE  VAVIQPPEYN  RFANSFGHIF
601 ACGYSAGYYS  YAWAEVLSTD  AYAAFEESDD  VAATGKRWFQ  EILAVGGSRS
651 AAESFKAFRG  REPSIDALLR  QSGFDNAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 483>:

```
m128.seq (partial)
1  ATGACTGACA  ACGCACTGCT  CCATTTGGGC  GAAGAACCCC  GTTTTGATCA
51  AATCAAAACC  GAAGACATCA  AACC CGCCCT  GCAAACCGCC  ATCGCCGAAG
101 CGCGCGAACA  AATCGCCGCC  ATCAAAGCCC  AAACGCACAC  CGGCTGGGCA
151 AACACTGTCTG  AACCCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201 GGGCGTGGTG  TCGCACCTCA  ACTGCGTCGC  CGACACGCCC  GAACTGCGCG
251 CCGTCTATAA  CGAACTGATG  CCCGAAATCA  CCGTCTTCTT  CACCGAAATC
301 GGACAAGACA  TCGAGCTGTA  CAACCGCTTC  AAAACCATCA  AAAATTCCCC
351 CGAATTCGAC  ACCCTCTCCC  CCGCAGAAA  AACCAGAACTC  AACCAC
1  TACGCCAGCG  AAAAAGTGGC  CGAAGCCAAA  TACGCGTTCA  GCGAAACCGA
51 wGTCAAAAAA  TAYTTCCCYG  TCGGCAAwGT  ATTAACCGGA  CTGTTCCGCC
101 AAmTCAAAAA  ACTmTACGGC  ATCGGATTTA  CCGAAAAAAC  yGTCCCGGTC
151 TGGCAGAAAG  ACGTGCGCTA  TtkTGAATTG  CAACAAAACG  GCGAAmCCAT
201 AGGCGGCGTT  TATATGGATT  TGTACGCACG  CGAAGGCAA  CGCGGCGGCG
251 CGTGGATGAA  CGACTACAAA  GGCCGCGGCC  GTTTTTCAGA  CGGCACGCTG
301 CAAYTGCCCA  CCGCCTACCT  CGTCTGCAAC  TTCGCCCCAC  CCGTCGGCGG
351 CAGGGAAGCC  CGCyTGAGCC  ACGACGAAAT  CCTCATCTC  TTCCAGGAAA
401 CCGGACACGG  GCTGCACCAC  CTGCTTACCC  AAGTGGACGA  ACTGGGCGTA
451 TCCGGCATCA  ACGGCGTAKA  ATGGGACGCG  GTCGAACTGC  CCAGCCAGTT
501 TATGAAAAAT  TTCGTTTGGG  AATACAATGT  CTTGGCAGAA  mTGTCAGCCC
551 ACGAAGAAAC  CGGcgTCCCC  yTGCCGAAAG  AACTCTTsGA  CAAAwTGCTC
601 GCCGCCAAAA  ACTTCCAAsg  CGGCATGTTC  yTsGTCCGGC  AAwtGGAGTT
651 CGCCCTCTTT  GATATGATGA  TTTACAGCGA  AGACGACGAA  GGCCGTCTGA
701 AAAACTGGCA  ACAGGTTTGA  GACAGCGTGC  GCAAAAAAGT  CGCCGTCTATC
751 CAGCCGCCCG  AATACAACCG  CTTCGCCTTG  AGCTTCGGCC  ACATCTTCGC
801 AGGCGGCTAT  TCCGCAGCTn  ATTACAGCTA  CGCGTGGGCG  GAAGTATTGA
851 GCGCGGACGC  ATACGCGGCC  TTTGAAGAAA  GCGACGATGT  CGCCGCCACA
901 GGCAACGCT  TTTGGCAGGA  AATCCTCGCC  GTCGGGGnAT  CGCGCAGCGG
951 nGCAGAAATCC  TTCAAAGCCT  TCCGCGGCCG  CGAACCGAGC  ATAGACGCAC
1001 TCTTGCGCCA  CAGCGGTTTC  GACAACGCGG  TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 484; ORF 128>:

```
m128.pep (partial)
1  MTDNALLHLG  EEPRFDQIKT  EDIKPALQTA  IAEAREQIAA  IKAQHTGWGA
51  NTVEPLTGIT  ERVGRIWGVV  SHLNCVADTP  ELRAVYNELM  PEITVFFTEI
101 GQDIELYNRF  KTIKNSPEFD  TLSPAQKTKL  NH

//
1  YASEKLREAK  YAFSETXVKK  YFPVGXVLNG  LFAQXKKLYG  IGFTEKTPV
51  WHKDVRVYXEL  QONGEXIGV  YMDLYAREGK  RGGAWMNDYK  GRRRFSDGTL
101 QLPTAYLCVN  FAPPVGGREA  RLSHDEILIL  FHETGHGLHH  LLTQVDELGV
151 SGINGVXWDA  VELPSQFMEN  FVWEYNVLAQ  XSAHEETGVP  LPKELXDKXL
201 AAKNFQXGMF  XVRQXEFALF  DMMIYSEDDE  GRLKNWQQVL  DSVRKKVAVI
251 QPPEYNRFAL  SFGHIFAGGY  SAAXYSYAWA  EVLSADAYAA  FEESDDVAAT
```

365

301 GKRFWQEILA VGXSRGAES FKAFRGREPS IDALLRHSGF DNAV*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQHTGTWANTVERLTGIT					
	: :					
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
g128.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA					
	:					
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130	140	150	160	170	180
g128.pep	TLSPAQKTKLDHDLRDFVLGAELPPERQAEALAKLTGEGAQLSAKFSQNVLDATDAFGIY					
	:					
m128	TLSPAQKTKLNH					
	130					
	//					
			340	350	360	
g128.pep	YAGEKLREAKYAFSETEVKKYFPVGKVLG					
	:					
m128	YASEKLREAKYAFSETXVKKYFPVGXVLNG					
			10	20	30	
	370	380	390	400	410	420
g128.pep	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWMNDYK					
	:					
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQONGEXIGGVYMDLYAREGKRGGAWMNDYK					
	40	50	60	70	80	90
	430	440	450	460	470	480
g128.pep	GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV					
	:					
m128	GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVDELGV					
	100	110	120	130	140	150
	490	500	510	520	530	540
g128.pep	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQGMF					
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF					
	160	170	180	190	200	210
	550	560	570	580	590	600
g128.pep	LVRQMEFALFDMMIYSESEDECRLLKNWQQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY					
	: :					
m128	XVRQXEFALFDMMIYSEDEDEGRLLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY					
	220	230	240	250	260	270
	610	620	630	640	650	660
g128.pep	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS					
	: :					
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRGAESFKAFRGREPS					
	280	290	300	310	320	330

366

	670	679
g128.pep	IDALLRQSGFDNAAX	
	: :	
m128	IDALLRHSGFDNAVX	
	340	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 485>:

```

a128.seq
1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTC
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCG GAACTGCGCG
251 CGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAATC AACACGATC
401 TGCGCGATTT CGTCCTCAGC GCGCGGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CCGCATTAC TTTGACGATG
551 CCGCACCCTG TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCGCGCCGCG CCAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCAT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTATTAAACG GACTGTTCCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCAGTCAG
1501 TTTATGGAAA ATTTCTGTTT GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATCTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCTG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 486; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWAA
51  NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```

367

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
 601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
 651 AAESFKAFRG REPSIDALLR HSGFDNAA*

m128/a128 66.0% identity in 677 aa overlap

	10	20	30	40	50	60
m128.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQHTGTWANTVEPLTGIT					
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQHTGTWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
m128.pep	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130					
m128.pep	TLSPAQKTKLNH-----					
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
				140	150	
m128.pep	-----YASEKLREAKYAFSETXVKKYFPVGX					
a128	ARRAKPYAEKDLAEVKAFARESGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
	160	170	180	190	200	210
m128.pep	VLNGLFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
	220	230	240	250	260	270
m128.pep	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
	280	290	300	310	320	330
m128.pep	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540
	340	350	360	370	380	390
m128.pep	XGMFVVRQXEFALFDMMIYSEDDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600

368

```

          400      410      420      430      440      450
m128.pep  AGGYSAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRG
          |||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128      AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
          610      620      630      640      650      660

          460      470
m128.pep  REPSIDALLRHSGFDNAVX
          ||||| ||||| ||||| |||||
a128      REPSIDALLRHSGFDNAAX
          670

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 487>:

```

g128-1.seq (partial)
1  ATGATTGACA  ACGCACTGCT  CCACTTGGGC  GAAGAACCCC  GTTTTAATCA
51  AATCAAAACC  GAAGACATCA  AACCCGCCGT  CCAAACCGCC  ATCGCCGAAG
101 CGCGCGGACA  AATCGCCGCC  GTCAAAGCGC  AAACGCACAC  CGGCTGGGCG
151 AACACCGTCG  AGCGTCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201 GGGCGTCTGT  TCCCATCTCA  ACTCCGTCGT  CGACACGCCC  GAACTGCGCG
251 CCGTCTATAA  CGAACTGATG  CCTGAAATCA  CCGTCTTCTT  CACCGAAATC
301 GGACAAGACA  TCGAACTGTA  CAACCGCTTC  AAAACCATCA  AAAATTCCCC
351 CGAATTTGCA  ACGCTTTCCC  CCGCACAAA  AACCAAGCTC  GATCACGACC
401 TGCGCGATTT  CGTATTGAGC  GCGCGGAAC  TGCGCGCCGA  ACGGCAGGCA
451 GAACTGGCAA  AACTGCAAAC  CGAAGGCGCG  CAACTTTCCG  CCAAATCTCT
501 CCAAACGTC  CTAGACGCGA  CCGACGCGTT  CGGCATTAC  TTTGACGATG
551 CCGCACCGCT  TGCGGCATT  CCGAAGACG  CGCTCGCCAT  GTTGCCGCC
601 GCGCGCAAAA  GCGAAGGCAA  AACAGGTTAC  AAAATCGGCT  TGCAGATTCC
651 GCACTACCTT  GCCGTTATCC  AATACGCCGG  CAACCGCGAA  CTGCGCGAAC
701 AAATCTACCG  CGCCTACGTT  ACCCGTGCCA  GCGAACTTTC  AAACGACGGC
751 AAATTCGACA  AACCGCCAAA  CTGCTCGGCT  TAAAAAATTA  CGCCGAATTG  TCGCTGGCAA
801 AACCGCCAAA  CTGCTCGGCT  TAAAAAATTA  CGCCGAATTG  TCGCTGGCAA
851 CCAAATGGC  GGACACGCCC  GAACAGGTTT  TAACTTCTCT  GCACGACCTC
901 GCGCGCGCG  CCAAACCTTA  CGCCGAAAA  GACCTCGCCG  AAGTCAAAGC
951 CTTCGCCCG  GAACACCTCG  GTCTCGCCGA  CCGCAGCCG  TGGGACTTGA
1001 GCTACGCCGG  CGAAAACTG  CGCGAAGCCA  AATACGCATT  CAGCGAAACC
1051 GAAGTCAAAA  AATACTTCCC  CGTCGGCAAA  GTTCTGGCAG  GCCTGTTCCG
1101 CCAATCAAAA  AAATCTACG  GCATCGGATT  CGCCGAAAA  ACCGTTCCCG
1151 TCTGGCACA  AGACGTGCGC  TATTTGAAT  TGCAACAAA  CGGCAAAACC
1201 ATCGCGGCG  TTTATATGGA  TTTGTACGCA  CGCGAAGGCA  AACGCGGCGG
1251 CGCGTGGATG  AACGACTACA  AAGGCGCGCG  CCGCTTTGCC  GACGGCACGC
1301 TGCAACTGCC  CACCGCCTAC  CTCGTCTGCA  ACTTCGCCCC  GCGCGTCGGC
1351 GGCAAGAAGA  CGCGTTAAG  CCACGACGAA  ATCCTCACCC  TCTTCCACGA
1401 AACC GGCCAC  GGA CTGCACC  ACCTGCTTAC  CCAAGTGGAC  GAACTGGGCG
1451 TGTCCGCAT  CAACGCGTA  AAA

```

This corresponds to the amino acid sequence <SEQ ID 488; ORF 128-1.ng>:

```

g128-1.pep (partial)
1  MIDNALLHLG  EEPRFNQIKT  EDIKPAVQTA  IAEARGQIAA  VKAQHTGW
51  NTVERLTGIT  ERVGRWGVV  SHLNSVVDTP  ELRAVYNELM  PEITVFFTEI
101 GQDIELYNRF  KTIKNSPEFA  TLSPAQKTKL  DHDLRDFVLS  GAELPPERQA
151 ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201 AAQSEKGTGY  KIGLQIPHYL  AVIQYAGNRE  LREQIYRAYV  TRASELSNDG
251 KFDNTANIDR  TLENALKTAK  LLGFKNYAEL  SLATKMADTP  EQVLNFLHDL
301 ARRAKPYAEK  DLAEVKAFAR  EHLGLADPQP  WDSL YAGEKL  REAKYAFSET
351 EVKKYFPVGK  VLAGLFAQIK  KLYGIGFAEK  TVPVVHKDVR  YFELQONGKT
401 IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFA  DGTLLQLPTAY  LVCNFAPPVG
451 GKEARLSHDE  ILTLFHETGH  GLHHLLTQVD  ELGVSGINGV  K

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 489>:

```

m128-1.seq
1  ATGACTGACA  ACGCACTGCT  CCATTTGGGC  GAAGAACCCC  GTTTTGATCA
51  AATCAAAACC  GAAGACATCA  AACCCGCCCT  GCAAACCGCC  ATCGCCGAAG
101 CGCGCGAACA  AATCGCCGCC  ATCAAAGCCC  AAACGCACAC  CGGCTGGGCA
151 AACACTGTCT  AACCCCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG

```

```

201 GGGCGTGGTG TCGCACTCA ACTCCGTGCG CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCTC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACCACGATC
401 TGCGCGATTT CGTCTCAGC GGCGCGGAAC TGCCGCCCCG ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAAC CGAAGGCGCG CAACTTTCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 ACACTACCTC GCCGTCAATC AATACGCGCA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCTACGTT ACCCGCGCCA GCGAATTTT AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAAC TAAGCGAATT TCGCTGGCAA
851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCGCGCCGCG CCAAACCCTA CGCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTTCG
1101 CCAAATCAAA AAACCTCTAC GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCTAC CTGCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGCAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTGCAACT GCCCAGCCAG
1501 TTTATGGAAT ATTTGCTTTG GGAATACAAT GTCTTGGCAC AAATGTGAGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAAATGC
1601 TCGCGGCCAA AAACCTTCCA CGCGGCATGT TCCTCGTCCG GCAAAATGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751 TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAACAG CTTTGGGAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

m128-1.pep.

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWAA
51 NTVPLTGIT ERVGRIVGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101 QQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFAPPVG
451 GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAFAEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAV*

```

m128-1/g128-1 94.5% identity in 491 aa overlap

```

          10      20      30      40      50      60
g128-1.pep MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90     100     110     120
g128-1.pep ERVGRIVGVVSHLNSVVDTPPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m128-1      ERVGRIVGVVSHLNSVADTPPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90     100     110     120

```


370

	130	140	150	160	170	180
g128-1.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
	190	200	210	220	230	240
g128-1.pep	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV					
m128-1	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV					
	190	200	210	220	230	240
	250	260	270	280	290	300
g128-1.pep	TRASELSNDGKFDNTANIDRTLLENALKTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL					
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL					
	250	260	270	280	290	300
	310	320	330	340	350	360
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDLSYAGEKLREAKYAFSETEVKKYFPVGK					
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPQWDLGYASEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
	370	380	390	400	410	420
g128-1.pep	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM					
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
	430	440	450	460	470	480
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
	490					
g128-1.pep	ELGVSGINGVK					
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 491>:

a128-1.seq

1	ATGACTGACA	ACGCACTGCT	CCATTGGGC	GAAGAACCCC	GTTTGTATCA
51	AATCAAAACC	GAAGACATCA	AACCCGCCCT	GCAAACCGCC	ATTGCCGAAG
101	CGCGCGAACA	AATCGCCGCC	ATCAAAGCCC	AAACGCACAC	CGGCTGGGCA
151	AACACTGTCTG	AACCCCTGAC	CGGCATCACC	GAACGCGTCG	GCAGGATTGG
201	GGGCGTGGTG	TCGCACCTCA	ACTCCGTCAC	CGACACGCCC	GAACGCGCG
251	CCGCCTACAA	TGAATTAATG	CCCGAAATTA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TCGAGCTGTA	CAACCGCTTC	AAAACCATCA	AAAACCTCCC
351	CGAGTTCGAC	ACCTCTCTCC	ACGCGCAAAA	AACCAAACTC	AACCACGATC
401	TGCGCGATTT	CGTCTCAGC	GGCGCGGAAC	TGCCGCCCGA	ACAGCAGGCA
451	GAATTGGCAA	AACTGCAAAAC	CGAAGGCGCG	CAACTTTCGG	CCAAATTCTC
501	CCAAAACGTC	CTAGACGCGA	CCGACGCGTT	CGGCATTTAC	TTTGACGATG
551	CCGCACCGCT	TGCCGCGATT	CCCGAAGACG	CGCTCGCCAT	GTTTGCCGCT
601	GCCGCGCAAA	GCGAAGGCAA	AACAGGCTAC	AAAATCGGTT	TGCAGATTCC
651	GCACTACCTC	GCCGTATCTC	AATACGCCGA	CAACCGCAAA	CTGCGCGAAC
701	AAATCTACCG	CGCTACGTT	ACCCGCGCCA	GCGAGCTTTC	AGACGACGGC
751	AAATTCGACA	ACACCGCCAA	CATCGACCGC	ACGCTCGAAA	ACGCCCTGCA
801	AACCGCCAAA	CTGCTCGGCT	TCAAAAATA	CGCCGAATTG	TCGCTGGCAA
851	CCAAAATGGC	GGACACCCCC	GAACAAGTTT	TAAACTTCCT	GCACGACCTC
901	GCCCGCGCG	CCAAACCCTA	CGCCGAAAAA	GACCTCGCCG	AAGTCAAAGC
951	CTTCGCCCGC	GAAAGCCTCG	GCCTCGCCGA	TTTGCAACCG	TGGGACTTGG
1001	GCTACGCCGG	CGAAAACTG	CGCGAAGCCA	AATACGCATT	CAGCGAAACC
1051	GAAATCAAAA	AATACTTCCC	CGTCGGCAAA	GTATTAAACG	GACTGTTTCG
1101	CCAAATCAAA	AAACTCTACG	GCATCGGATT	TACCGAAAAA	ACCGTCCCCG
1151	TCTGGCACA	AGACGTGCGC	TATTTTGAAT	TGCAACAAAA	CGGCGAAACC
1201	ATAGGCGGCG	TTTATATGGA	TTGTACGCA	CGCGAAGGCA	AACGCGGCGG

```

1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACCTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCT
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCGCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTGWA
51 NTVEPLTGIT ERVGRINGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQTKKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLPQ WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVVHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*

```

m128-1/a128-1 97.8% identity in 677 aa overlap

a128-1.pep	10	20	30	40	50	60
m128-1	10	20	30	40	50	60
a128-1.pep	70	80	90	100	110	120
m128-1	70	80	90	100	110	120
a128-1.pep	130	140	150	160	170	180
m128-1	130	140	150	160	170	180
a128-1.pep	190	200	210	220	230	240
m128-1	190	200	210	220	230	240
a128-1.pep	250	260	270	280	290	300
m128-1	250	260	270	280	290	300
a128-1.pep	310	320	330	340	350	360
m128-1	310	320	330	340	350	360

```

m128-1      ARRAKPYAEKDLAEVKAFARESLNLADLPWDLGYASEKLREAKYAFSETEVKKYFPVGK
              310      320      330      340      350      360

              370      380      390      400      410      420
a128-1.pep  VLNLGFAQIKKLYGIGFTEKTVPVVHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
              |||||||
m128-1      VLNLGFAQIKKLYGIGFTEKTVPVVHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
              370      380      390      400      410      420

              430      440      450      460      470      480
a128-1.pep  NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHLHLLTQVD
              |||||||
m128-1      NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHLHLLTQVD
              430      440      450      460      470      480

              490      500      510      520      530      540
a128-1.pep  ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
              |||||||
m128-1      ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
              490      500      510      520      530      540

              550      560      570      580      590      600
a128-1.pep  RGMFLVRQMEFALFDMMIYSEDDEGRLLKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF
              |||||||
m128-1      RGMFLVRQMEFALFDMMIYSEDDEGRLLKNWQQVLDVSRKKAQVIQPPPEYNRFALSFGHIF
              550      560      570      580      590      600

              610      620      630      640      650      660
a128-1.pep  AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
              |||||||
m128-1      AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
              610      620      630      640      650      660

              670      679
a128-1.pep  REPSIDALLRHSGFDNAAX
              |||||||
m128-1      REPSIDALLRHSGFDNAVX
              670

```

a128-1/ P44573

sp|P44573|OPDA_HAEIN OLIGOPEPTIDASE A >gi|1075082|pir||C64055 oligopeptidase A (prlc) homolog
 - Haemophilus influenzae (strain Rd KW20)
 >gi|1573174 (U32706) oligopeptidase A (prlc) [Haemophilus influenzae Rd] Length = 681
 Score = 591 bits (1507), Expect = e-168
 Identities = 309/677 (45%), Positives = 415/677 (60%), Gaps = 4/677 (0%)

```

Query: 4  NALLHLGEEPRFDQIKTEDIKPALQTXXXXXXXXXXXXXXTHGTWANTVEPLTGITERV 63
          N LL++ P F QIK E I+PA++ H W N + PLT +R+
Sbjct: 5  NPLLNIQGLPPFSQIKPEHIRPAVEKLIQDCRNTIEQVLKQPHFTWENFILPLTETNDRL 64

Query: 64 GRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFDTLS 123
          R W VSHLNSV ++ ELR AY +P ++ + T +GQ LYN + +KNS EF S
Sbjct: 65 NRAWSPVSHLNSVKNSTELREAYQTCPLLLSEYSTWVGQHKGLYNAYLALKNSAEFADYS 124

Query: 124 HAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIYFDD 183
          AQK + + LRDF LSG L E+Q ++ ++L+++FS NVLDAT + ++
Sbjct: 125 IAQKKAIEINLRDFELSGIGLSEEKQORYGEIVARLSELNSQFSNNVLDATMGWEKLIEN 184

Query: 184 AAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLRQIYRAYVTRA 243
          A LAG+PE AL +A+S+G GY+ L+IP YL V+ Y +NR LRE++YRAY TRA
Sbjct: 185 EAELAGLPESALQAAQQAESKGLKGYRFTLEIPSYLPVMTYCENRALREEMYRAYATRA 244

Query: 244 SELSDD-GKFDNTANIDRTLENALQTAKLGLGFKNYAELSLATKMAADTPEQVLNFLHDLAR 302
          SE + GK+DN+ ++ L ++ AKLLGF Y ELSLATKMA+ P+QVL+FL LA
Sbjct: 245 SEQGPNAKGWDNSKVMEEILTLRVELAKLLGFNTYTELSLATKMAENPQQVLDFLDLHAE 304

Query: 303 RAKPYAEKDLAEVKAFARESLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGKVL 362
          RAKP EK+L E+K + + G+ +L PWD+G+ EK ++ YA ++ E++ YFP +V+
Sbjct: 305 RAKPQGEKELQELKGYCEKEFGVTELAPWDIGFYSEKQKQHLAYAINDEELRPYPENRVI 364

```

Query: 363 NGLFAQIKKLYGIGFTE-KTVPVWHKDVRYFEL-QQNGETIGGVYMDLYAREGKRGGAWM 420
 +GLF IK+++ I E K V WHKDVR+F+L +N + G Y+DLYARE KRGGAWM
 Sbjct: 365 SGLFELIKRIFNIRAVRKGVDTHKDVRFDDLIDENDQLRGSFYLDLYAREHKRGGAWM 424

Query: 421 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEIXXXXXXXXXXXXXXXXXXQVD 480
 +D GR+R DG+++ P AYL CNF P+G K A +H+E+ Q+D
 Sbjct: 425 DDCIGRKRKLDGSIETPVAYLTCNFPNAPIGNKPALFTHNEVTTLFHEFGHGIHMLTQID 484

Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540
 V+GINGV WDAVELPSQFMEN+ WE LA +S H ETG PLPKE ++L AKNFQ
 Sbjct: 485 VSDVAGINGVPWDVELPSQFMENWCWEEELAFISGHYETGEPLPKEKLTQLLKAKNFQ 544

Query: 541 RGMFLVRQMEFALFDMMIYSEDDEGRKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF 600
 MF++RQ+EF +FD ++ D + L SV+ +VAV++ ++ R +SF HIF
 Sbjct: 545 AAMFILRQLEFGIFDRLHHTFDAEKTNQILDTLKSVKSQVAVIKGVDWARAPHSFSHF 604

Query: 601 XXXXXXXXXXXXWAEVLSADAYAAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659
 WAEVLSADAY+ FEE TKG F EIL GGS E FK FR
 Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEGIFNPITGKSFLDELTRGGSEEPMELEFKRFR 664

Query: 660 GREPSIDALLRHSGFDN 676
 GREP +DALLRH G N
 Sbjct: 665 GREPQLDALLRHKGIMN 681

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 493>:

g129.seq
 1 ATGCTTTCAC CTCCTCGGCG TAAACGGCG GCACATCAAT CAAGCCGTCT
 51 TTCATTGCG TCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT
 101 ATCGTGCAGC ATCCAGCCCA AATCGCGGTT TGCCTCGCTT TCCGATAACG
 151 CCGACGGCGG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTTGCCCTT
 201 TCAGGCTGCG GGCATAGGGG CGGAACAGGC GCGGTCGAA TCCTGTTCA
 251 TCCGGACAAA CGCGTTGGCA GTCGGAAAAAT CCGGCCGCC GTGTCAAATA
 301 ATGCGTTACT TTGGCCGGT CTTGTCCTTT GTAAGCGGCG GTCTTTTTTT
 351 GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC
 401 AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCGG ATGCCGCCCG
 451 ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TTAGGCCGGT
 501 AACTTGA

This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:

g129.pep
 1 MLSPPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT
 51 PTAAAVHPYP RFRHLPFOAA GIGAEQAAVE SCFIRTNALA VGKSGRPCQI
 101 MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP
 151 TYRAGFCLSD LAAFRPVT*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 495>:

m129.seq (partial)
 1 ..TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
 51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
 101 GAAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGCTCTTG
 151 TTCTTTGTAA GTGGTGGTCT TTTTTTGCGC GTTATCCCCA TCTGTTTGAG
 201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT
 251 GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTGAGCCGG TTTTGCCTA
 301 TCCGATTTGA CGGCATTTAG ACCGGTAACT TGA

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

m129.pep (partial)
 1 ..YLRPHYLPFO AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGRVL
 51 FFVSGGLFLR VIPICLSAQ MVAQVSKCL AISCRXASGC CPTYXAGFCL
 101 SDLTAFRPVT *

Computer analysis of this amino acid sequence gave the following results:

m129/g129

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 497>:

This corresponds to the amino acid sequence <SEQ ID 498; ORF 129.a>:

m129/a129 98.2% identity in 110 aa overlap

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 499>:

g130.seq					
1	ATGAAACAAC	TCCGCGACAA	CAAAGCCCAA	GGCTCTGCAC	TGTTTACCCT
51	TGTGAGCGGT	ATCGTTATTG	TTATTGCAGT	CCTTTATTTT	CTGATTAAAGC
101	TGGCGGGCAG	TGGATCGTTT	GGCGATGTCG	ATGCCACTAC	GGAAGCGGCA
151	ACGCAGACCC	GCATCCAGCC	TGTCGGACAA	TTGACGATGG	GTGACGGCAT
201	CCCCGTGGC	GAAAGCCCAAG	CGCAACAGAT	TTTCGGCAAA	ATCTGTATCC

375

```

251 AATGCCACGC GCGGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGACT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGCAG
401 ACCTGACCGA TCAGGAATC AAACGGGCGA TTACCTACAT GGCGAATAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTGAGGTG TGCCACGGCG GTTCGATTCC
651 CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGAC AAACATGCCC TTGAAGGCTT TAACGCGATG
751 CCGGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 500; ORF 130.ng>:

```

g130.pep
1  MKQLRDNKAQ  GSALFTLVSG  IVIVIAVLYF  LIKLAGSGSF  GDVDATTEAA
51  TQTRIQPVGQ  LTMGDGIPVG  ERQGEQIFGK  ICIQCHAADS  NVPNAPKLEH
101 NGDWAPRIAQ  GFDTLFQHAL  NGFNAMPAKG  GAADLTQDEL  KRAITYMANK
151 SGGSFNPDE  AAPADNAASG  TASAPADSAA  PAEAKAEDKG  AAPAVGVVDG
201 KKVFEATCQV  CHGGSIPGIP  GIGKKDDWAP  RIKKGKETHL  KHALEGFNAM
251 PAKGNGAGLS  DDEVKAAVDY  MANQSGAKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 501>:

```

m130.seq (partial)
1  ..GGCGAACAGA TTTTCGGCAA AATCTGTATC CAATGCCACG CGGCGGACAG
51  CAATGTGCCG AACGCTCCGA AACTGGAACA CAACGGCGAT TrGGCACCGC
101 GTATCGgCAA GGCTTCGATA CCTTGTTCCA ACACGCGCTG AACGGCTTTA
151 ACGCCATGCC TGCAAAGGC GGTGCGGCAG ACCTGACCGA TCAGGAACCTT
201 AAACGGGCGA TTACTTACAT GGCGAACAAA AGCGGCGGTT CTTTCCCGAA
251 TCCTGATGAG GCTGCGCCTG CCGACAATGC CGCTTCAGGA ACAGCTTCTG
301 CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG CGAAGGCAGA AGACAAGGGT
351 GCGGCACCCC TGCGGTCCGG GTTGACGGTA AAAAAAGTCTT CGAAGCAACC
401 TGTCAGGTGT GCCACGGCGG TTCGATTCCC GGTATTCCCG GCATAGGCAA
451 AAAAGACGAT TGGGCACCGC GTATCAAAAA AGGCAAAGAA ACCTTGCACA
501 AACACGCCCT TGAAGGCTTT AACGCGATGC CTGCCAAArG CGgCAATGCA
551 GGTTTGAGCG ATGACGAAgT CAAAGCGGCT GTTGACTATA TGGCAAACCA
601 ATCCGGTGCA AAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 502; ORF 130>:

```

m130.pep (partial)
1  ..GEQIFGKICI QCHAADSNVP NAPKLEHNGD XAPRIQGFDT LFQHALNGFN
51  AMPAKGGAAD LTDQELKRAI TYMANKSGGS FPNPDEAAPA DNAASGTASA
101 PADSAPAEA KAEDKGAAPA VGVDGKKVFE ATCQVCHGGS IPGIPGIGKK
151 DDWAPRIKKG KETLHKHALE GFNAMPAXKG NAGLSDDDEVK AAVDYMANQS
201 GAKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 130 shows 98.1% identity over a 206 aa overlap with a predicted ORF (ORF 130.ng) from *N. gonorrhoeae*:

m130/g130

```

m130.pep
                                     10      20      30
                                     GEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     |||
g130      DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     50      60      70      80      90     100

                                     40      50      60      70      80      89
m130.pep      XAPRI-QGFDTLFQHALNGFNAMPAKGGAADLTQDELKRAITYMANKSGGSFPNPDEAAP
                                     |||
g130      WAPRIAQGFDTLFQHALNGFNAMPAKGGAADLTQDELKRAITYMANKSGGSFPNPDEAAP
                                     110     120     130     140     150     160

```

376

```

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAPAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||
g130       ADNAASGTASAPADSAAPAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDDEVKAAVDYMANQSGAKFX
          |||
g130       KKDDWAPRIKKGKETLHKHALEGFNAMPKGGNAGLSDDDEVKAAVDYMANQSGAKFX
          230      240      250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 503>:

```

a130.seq
1   ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCCT
51  TGTGAGCGGT ATCGTTATTG TTATTGCACT CCTTTATTTC CTGATTAAGC
101 TGGCGGGCAG CGGCTCGTTC GGCGATGTCG ATGCCACTAC GGAAGCAGCA
151 ACGCAGACCC GTATCCAGCC TGTCGGACAA TTGACGATGG GCGACGGCAT
201 CCCCCTCGGC GAACGCCAAG GCGAACAGAT TTTCGGCAA ATCTGTATCC
251 AATGCCACGC GCGGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGATT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGTAG
401 ACCTGACCGA TCAGGAATC AAACGGGCGA TTAATTACAT GGCGAACAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CCGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAGA AACCTTGCAC AAACACGCCC TTGAAGGCTT TAACGCGATG
751 CCTGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 504; ORF 130.a>:

```

a130.pep
1   MKQLRDNKAQ GSALFTLVSG IVIVIAVLVF LIKLAGSGSF GDVDATTEAA
51  TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFOHAL NGFNAMPAKG GAVDLTDQEL KRAITYMANK
151 SGGSFNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAPAVGVDG
201 KKVFEATCQV CHGGSIPGIP GIGKKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

m130/a130 97.6% identity in 206 aa overlap

```

          10      20      30
m130.pep  GEQIFGKICIQCHAADSNVPNAPKLEHNGD
          |||
a130       DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
          50      60      70      80      90      100

          40      50      60      70      80      89
m130.pep  XAPRI-QGFDTLFOHALNGFNAMPAKGGAADLTDQELKRAITYMANKSGGSFNPDEAAP
          |||
a130       WAPRIAQGFDTLFOHALNGFNAMPAKGGAADLTDQELKRAITYMANKSGGSFNPDEAAP
          110      120      130      140      150      160

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAPAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||
a130       ADNAASGTASAPADSAAPAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDDEVKAAVDYMANQSGAKFX

```

|||||
a130 KKDDWAPRIKKGKETLHKHALEGFNAMPAGGNAGLSDDEVKAAVDYMANQSGAKFX
230 240 250 260 270 280

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 505>:

g132.seq
1 ATGGAAGCCT TCAAAACCCT AATTTGGATT ATTAATATTA TTTCCGCTTT
51 GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151 GCCGGCAACG CCAACTTcct CAGccGCTCG AccGccGTTG CAGCAACAtt
201 tttcttTGca acctgcAtgg gctatggTgt atattcacac CCACACGACA
251 AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
301 ACCcgTAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTAG CAGCAGAAAT
351 AACagtTTTT CAAATgccga catGgtga

This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:

g132.pep
1 MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
51 AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
101 TRKQYRTFCP CSSAAEITVF QMPTW*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 507>:

m132.seq (partial)
1 ATGGAACCCT TCAAAACCTT AATTTGGATT GTTAATTTAA TTTCCGCTTT
51 GGCCGTCCTC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACTTT CGGA...

This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:

m132.pep (partial)
1 MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG...

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng) from *N. gonorrhoeae*:

m132/g132

	10	20	30	
m132.pep	MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG			
	: : :			
g132	MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSGSGSAQGVFGSAGNANFLSRS			
	10	20	30	40 50 60

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 509>:

a132.seq
1 ATGGAAGCCT TCAAAACCCT AATTTGGATT GTTAATATAA TTTCCGCTTT
51 GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151 GCCGGCAACG CTAACCTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
201 TTTCTTTGCA ACCTGCATGg GCTATGGTGT ATATTCACAC CCACACGACA
251 AAACACGGTT TGGACTTCAG CAACGTACAA CAACTCAGC AAGCACCCAA
301 ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTAG CAGCAGAAAT
351 AACAGTTTTT CAAATGCCGA CATGGTGA

This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:

a132.pep
1 MEAFKTLIWI VNIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
51 AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
101 TRKQYRTFCP CSSAAEITVF QMPTW*

m132/a132 92.1% identity in 38 aa overlap

378

```

      10      20      30
m132.pep  MEPFKTLIWIVNLISALAVFVLVLLQHGKGADAGATFG
           || |||||:|||||:|||||:|||||:|||||:|||||
a132      MEAFKTLIWIVNIISALAVIVLVLLQHGKGADAGATFGSGSGSAQGVFGSAGNANFLSRS
           10      20      30      40      50      60

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 511>:

```

g134.seq
1  ATGTCCCAAG AAATCCTCGA CCAAGTGCGC CGCCGCCGCA CGTTTGCCAT
51  CATCTCCCAC CCCGATGCGG GTAAAACCCAC GCTGACCGAA AAAGTGTCTGC
101 TGTTCCTCGG CGCGATTCAA AGCGCAGGCA CGGTGAAAGG TAAGAAAACC
151 GGCAAATTCG CCACCTCCGA CTGGATGGAC ATCGAGAAGC AGCGCGGCAT
201 TTCCGTGGCA TCAAGCGTGA TGCAGTTCGA CTACAAAGAC CACACCGTCA
251 ACCTCTTGGA CACGCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTAACCG CAGTGGACAG CGCCTTGATG GTCATCGACG CGGCAAAAGG
351 CGTGGAAGCG CAAACCATCA AACTCTTGAA CGTCTGCCGC CTGCGCGATA
401 CGCCGATTGT TACCTTCATG AACAAATACG ACCGCGAAGT GCGCGATTCT
451 TTGGAACCTT TGGACGAAGT GGAAGACATC CTGCAAATCC GCTGCGCGCC
501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA
551 TCCTGAACGA CGAAATCTAT CTCTTTGAAG CGGGCGGCGA ACGCCTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATAAAC AATCCCGAAT TGGAACAACG
651 CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCCAA CGAATTAAAT CTCGacgaAT TTCTCGcgg CGAACTCACG
751 CCAGTGTCTT TCGGCTCTGC GATTAACAAC TTCGGCATTG AGGAAATCCT
801 CAATTCATTG ATTGACTGGG CACCCGCACC GAAACCGCGC GACGCGACCA
851 TGCGCATGGT CGGGCCGGAC GAGCCGAAAT TTTCCGGATT TATCTTTAAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATCG CCTTCTTGCG
951 CGTCTGCTCC GGTAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTAG TAACCTTCAT GTCGCACGAC
1051 CGCGAAGTGG CGGAAGAAGC CTACGCCGGC GACATCATCG GCATCCCGAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACCTG
1151 CGTTTACCGG CATCCCATTC TTCGCGCCCG AACTGTTCCG CAGCGTCCGC
1201 ATCAAAAACC CGCTGAAAAT CAAACAACG CAAAAAGGTT TGCAACAAC
1251 CGGCGAAGAA GGTGCGGTTC AAGTATTCAA ACCGATGAGC GGCGCGGATT
1301 TGATTTTGGG TGCGGTTCGG GTGTTGCACT TTGAAGTCGT AACCTCACGC
1351 CTCGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAATG GCGGAATTG
1451 AAAAAGCCAA CGCAGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTTGGGG TTGACGCAAG AACGCTGGCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAACA TTCGGTCAAA CTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 512; ORF 134.ng>:

```

g134.pep
1  MSQEILDQVR RRTFAIISH PDAGKTLTE KLLLFSGAIQ SAGTVKGKKT
51  GKFSATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSSEDYR
101 VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151 LELLDEVEDI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGIN NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFILAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATMRMVGPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELAEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGVAV FDSASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 513>:

```

m134.seq
1  ATGTCCCAAG AAATCCTCGA CCAAGTGCGC CGCCGCCGCA CGTTTGCCAT
51  CATCTCCCAC CCTGACGCG GTTAAACCCAC GTTGACTGAA AAAGTGTCTGC
101 TGTTCCTCGG CGCGATTCAA AGCGCGGGTA CGGTAAAAGG CAAGAAAACC
151 GGCAAATTCG CCACTTCCGA CTGGATGGAA ATCGAGAAGC AGCGCGGCAT

```

```

201  TTCCGTGGCA TCAAGTGTGA TGCAGTTCGA TTACAAAGAC CACACCGTCA
251  ACCTCTTGGA CACGCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301  GTTTTAACCG CCGTGGACAG CGCATTAAATG GTCATCGACG CGGCAAAAGG
351  CGTGGAAGCG CAAACCATCA AGCTCTTAAA CGTCTGCCGC CTGCGCGATA
401  CACCGATTGT TACGTTTATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451  CTGGAACCTT TGGACGAAGT GGAAAACATT TAAAAAATCC GCTGCGCGCC
501  CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA
551  TCCTGAACGA TGAAATTTAT CTCTTTGAAG CTGGCGGCGA ACGCCTGCCG
601  CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCTGAAT TGGAACAACG
651  CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701  CGGCTTCCAA CGAGTTTAAAT CTCGACGAAT TCCTCGCCGG CGAACTCACG
751  CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGGTATTC AGGAAATCCT
801  CAATTCATTG ATTGACTGGG CGCCCGCGCC GAAACCGCGC GACGCGACCG
851  TACGTATGGT CGAGCCGAC GAGCCGAAGT TTTCGGATT TATCTTCAAA
901  ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG
951  CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TTACCTTCAT GTCGCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GCATCCCGAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAATCG
1151 CGTTACCCGG CATCCCATTG TTCGCACCCG AACTGTTCCG CAGCGTACGC
1201 ATCAAAAACC CGCTGAAAAT CAAACAACG CAAAAAGGCT TGCAACAGCT
1251 CGGCGAAGAA GGCGCGGTGC AGGTGTTCAA ACCGATGAGC GGCGCGGATT
1301 TGATTTTGGG CGCGGTGCGC GTGTGTCAGT TTGAAGTCGT TACCTCGCGC
1351 CTCGCCAACG AATACGGCGT AGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAATG GCTGAATTTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTGCGG CTCACGCAAG AACGTTGGCC
1551 GGACATCGTG TTCCACGAAA CACGCGAACA TTCGGTCAAA CTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```

m134.pep
  1  MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
 51  GKFATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101  VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151  LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201  HEFDIIKGID NPELEQRFP L EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251  PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301  IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351  RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401  IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVTSR
451  LANEYGVAV FDSASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501  YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng) from *N. gonorrhoeae*:

```

m134/g134
      10      20      30      40      50      60
m134.pep  MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
          |||
g134      MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
          |||
      70      80      90     100     110     120
m134.pep  IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAKGVEA
          |||
g134      IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAKGVEA
          |||
      130     140     150     160     170     180

```

380

m134 . pep	QTIKLLNVCLRDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWPIGMGKNFKG
g134	QTIKLLNVCLRDTPIVTFMKNKYDREVRDSLELLDEVEDILQIRCAPVTWPIGMGKNFKG
	130 140 150 160 170 180
m134 . pep	VYHILNDEIYLF EAGGERLPHEFDI IKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
g134	VYHILNDEIYLF EAGGERLPHEFDI IKGINNPELEQRFPLEIQQLRDEIELVQAASNEFN
	190 200 210 220 230 240
m134 . pep	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPDEPKFSGFIFK
g134	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATMRMVGPDPEPKFSGFIFK
	250 260 270 280 290 300
m134 . pep	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFM SHDRELVEEAYAG
g134	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFM SHDRELAE EAYAG
	310 320 330 340 350 360
m134 . pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE
g134	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE
	370 380 390 400 410 420
m134 . pep	GAVQVFKPMMSGADLILGAVGVLFQFEVVT SRLANEYGV EAVFDSASIWSARWVSCDDKKKL
g134	GAVQVFKPMMSGADLILGAVGVLFQFEVVT SRLANEYGV EAVFDSASIWSARWVSCDDKKKL
	430 440 450 460 470 480
m134 . pep	AEFEKANAGNL AIDAGGNLAY LAPNRVNLGLTQERWPDIVFHETREHSVKLX
g134	AEFEKANAGNL AIDAGGNLAY LAPNRVNLGLTQERWPDIVFHETREHSVKLX
	490 500 510 520 530

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 515>:

```

a134 . seq
1  ATGTCCCAAG AAATCCTCGA CCAAGTGC GC CGCCGCCGCA CGTTTGCCAT
51  CATCTCCAC  CCTGACGCAG GTAAAACCAC GTTGA CTGAA AAAC TCTTGC
101 TGTTTT CAGG TGCGATTCAA AGCGCGGGTA CGGTAAAAGG CAAGAAAACC
151 GGCAAATTCG CCACCTCCGA CTGGATGGAC ATCGAGAAGC AGCGCGGCAT
201 TTCCGTGGCA TCAAGCGTGA TGCAGTTCGA CTATAAGAC CACACCGTCA
251 ACCTTTTGG A CACGCCGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTGACCG CCGTCGATAG TGCCTTGATG GTCATCGACG CGGCAAAAGG
351 CGTGGAAGCG CAAACCATCA AACTCTTGAA CGTCTGCCGC CTGCGCAATA
401 CGCCGATTGT TACGTT CATG AACAAATACG ACCCGGAAGT GCGCGATTCC
451 CTGGAATTGC TGGACGAAGT GGAACATC CTGCAATCC GCTGCGCGCC
501 CGTAACCTGG CCGATCGGCA TGGGCAAAAA CTTC AAAGGC GTGTACCACA
551 TCCTGAACGA CGAAATCTAT CTCTTGAAG CGGGCGGCGA ACGCTTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCCGAAT TGGAACAACG
651 CTTCCGTTA GAAATACAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCAA CGAGTTCAAT CTCGACGAAT TCCTGCGCGG CGAACTCACG
751 CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGTATT C AGGAAATCCT
801 CAATTCATTG ATTGAATGGG CGCCCGCGCC GAAACCACGC GATGCGACCG
851 TGCGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGATT TATCTTCAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG

```

```

951 CGCTGTGCTCC GGCAAATTCG AGCGCGGCAT GAAAATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TAACCTTCAT GTCCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GTATCCCAAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGA
1151 CGTTTACCGG CATCCCATTG TTCGCGCCCG AACTGTTCCG CAGCGTTCGC
1201 ATCAAAAACC CGCTGAAAAT CAAGCAACTG CAAAAAGGTT TGCAACAGCT
1251 TGGCGAAGAA GGTGCGGTGC AGGTGTTCAA ACCAATGAGC GGC GCGGATT
1301 TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351 CTTGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAACG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAAGT GCGGAATTTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCGGGCGG CAACCTCGCC
1501 TACCTCGCCC CTAACCGCGT GAATCTGGGA CTCACGCAAG AACGCTGGCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAGCA TTCGGTCAA CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 516; ORF 134.a>:

```

a134.pep
  1 MSQEILDQVR RRRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
  51 GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
 101 VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRNTPIVTFM NKYDREVRDS
 151 LELLDEVENI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
 201 HEFDIIKID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFELAGELT
 251 PVFFGSAINN FGIQEILNSL IEWAPAPKPR DATVRMVEPD EPKFSGFIFK
 301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
 351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLTFTGIPF FAPELFRSVR
 401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
 451 LANEYGVAV FDNASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA
 501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

m134/a134 98.9% identity in 531 aa overlap

	10	20	30	40	50	60
m134.pep	MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME					
a134	MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD					
	10	20	30	40	50	60
m134.pep	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA					
a134	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA					
	70	80	90	100	110	120
m134.pep	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA					
a134	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA					
	70	80	90	100	110	120
m134.pep	QTIKLLNVCRRLRDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWP					
a134	QTIKLLNVCRRLRNTPIVTFMKNKYDREVRDSLELLDEVENILQIRCAPVTWP					
	130	140	150	160	170	180
m134.pep	QTIKLLNVCRRLRDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWP					
a134	QTIKLLNVCRRLRNTPIVTFMKNKYDREVRDSLELLDEVENILQIRCAPVTWP					
	130	140	150	160	170	180
m134.pep	VYHILNDEIYLFEEAGGERLPHEFDIIKIDNPELEQRFPLEIQQLRDEIELVQAASNEFN					
a134	VYHILNDEIYLFEEAGGERLPHEFDIIKIDNPELEQRFPLEIQQLRDEIELVQAASNEFN					
	190	200	210	220	230	240
m134.pep	VYHILNDEIYLFEEAGGERLPHEFDIIKIDNPELEQRFPLEIQQLRDEIELVQAASNEFN					
a134	VYHILNDEIYLFEEAGGERLPHEFDIIKIDNPELEQRFPLEIQQLRDEIELVQAASNEFN					
	190	200	210	220	230	240
m134.pep	LDEFELAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK					
a134	LDEFELAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK					
	250	260	270	280	290	300
m134.pep	LDEFELAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK					
a134	LDEFELAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK					
	250	260	270	280	290	300
m134.pep	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG					
a134	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG					
	310	320	330	340	350	360
m134.pep	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG					
a134	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG					
	310	320	330	340	350	360

382

	370	380	390	400	410	420
m134.pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE					
	:					
a134	DIIGIPNHGNIQIGDSFSEGEQLTFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE					
	370	380	390	400	410	420
	430	440	450	460	470	480
m134.pep	GAVQVFKPMSGADLILGAVGVLFQFEVVTSLRANEYGVAVFDSASIWSARWVSCDDKKKL					
	:					
a134	GAVQVFKPMSGADLILGAVGVLFQFEVVTSLRANEYGVAVFDNASIWSARWVSCDDKKKL					
	430	440	450	460	470	480
	490	500	510	520	530	
m134.pep	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
a134	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
	490	500	510	520	530	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 517>:

g135.seq

```

1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCG
51  TTCGGAcgGC AGCCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCTG CAGGGTTCGG CGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GTGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCGATTCCC ATCATCAATG
401 AAAACGACAC GGTTTCGGTT GAGGAGTTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 CGGGGCGGCT CGGGTTCGGC AAACGGCACG GCGGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCACTCAA ACCCGATTCA TTGGCCGAAG CCGCCGAACA TCAGGCGGAC
751 GGCTCGTTT TCGTcccCcg tgCCAAAGGT TTGCGGACAC AGAAGCAATG
801 GctggCGTTC TATTcggaaa gcggGGgcag cgttTatgtg gacgaaagtg
851 cggaacacgc tTgtccgaa caagggaaa cctgCTGA

```

This corresponds to the amino acid sequence <SEQ ID 518; ORF 135.ng>:

g135.pep

```

1  MKYKRIVFKV GTSSITRSDG SLSRGKIQTI TRQLAALHHA GHELVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQORRAIP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIEM
201 AGGSGSANGT GGMLTKIKAA TIAAESGVPV YICSSLKPDS LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESGGSVYV DESAEHALSE QGKAC*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 519>:

m135.seq

```

1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCTGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCGG CAGGGTTCGG TGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACTG TCTTCAGACG GCATCGTGTC CGCGCAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTTCGGTT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT

```

```

501 GACCGACATA GACGGTCTTT ACACGGGCAA CCCGAACAGC AATCCCAGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AAACGGCACG GCGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CTGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCCGAAA GCCGGGGCAG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTTGTCCGAA CAGGGGAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGGAATCG AAGGGCATT TCCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCCC TGGGCAAAGG GCGCGTCTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATCGCGTA AGGCGAAAGG CGTGTTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG

```

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

```

m135.pep
  1 MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TCQLAALHHA GHVLVLVSSG
 51 AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAPV IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGPNPS NPDAVRLDKI EHINHEI IEM
201 AGSGSGSANGT GGMLTKIKAA TIAESGVPV YICSSLKPDA LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRGSVYV DEGAEHALSE QGKSLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAEDLL KSRKAKGVFI
351 HRDDWISITP EIRLLLTEF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng) from *N. gonorrhoeae*:

```

m135/g135
      10      20      30      40      50      60
m135.pep  MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      MKYKRIVFKVGTSSITRSDGSLSRGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG
          10      20      30      40      50      60

      70      80      90     100     110     120
m135.pep  FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL
          70      80      90     100     110     120

      130     140     150     160     170     180
m135.pep  SVLLQRRAPVIINENDTVSVEELKIGDNDTLAQVAAMIQADLLVLLTDIDGLYTGPNPS
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      SVLLQRRAPVIINENDTVSVEELKIGDNDTLAQVAAMIQADLLVLLTDIDGLYTGPNPS
          130     140     150     160     170     180

      190     200     210     220     230     240
m135.pep  NPDAVRLDKIEHINHEI IEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPDA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      NPDAVRLDKIEHINHEI IEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPD
          190     200     210     220     230     240

      250     260     270     280     290     300
m135.pep  LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRGSVYVDEGAEHALSEQGKSLMSGI
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESGSVYVDESAEHALSEQGKACX
          250     260     270     280     290

      310     320     330     340     350     360

```

m135.pep AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAAEDLLKSRKAKGVFIHRDDWISITP

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 521>:

```
a135.seq
1   ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTCGGG CAGGGTTTCGG TGCCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CCGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTCGGT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CCGGTTCGGC AAACGGCACA GGCGGTATGC TGAATAAAAT
651 CAAAGCGGCG ACGATTGCGA CCGAGTCCGG CGTACCGGTC TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CGGCAGATAA TCAGCGCGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCGAAA GCAGGGGCGG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTGTCCGAA CAGGGAAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGAATCG AAGGGCATT TTCCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCTT TGGGCAAAGG GCGAGTCTG TCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATTGCGTA AGGCGAAAGG CGTGTTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTAG
```

This corresponds to the amino acid sequence <SEQ ID 522; ORF 135.a>:

```
a135.pep
1   MKYKRIVFKV GTSSITHSDG SLRSGKIQTI TRQLAALHHA GHELVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRRAVP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGlyTGNPNS NPDAVRLDKI EHINHEIEM
201 AGGSGSANGT GGMLTKIKAA TIATESGVVP YICSSLKPDA LAEADNQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRRGVYV DEGAEHALSE QGKSLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAEDLL KLRKAKGVFI
351 HRDDWISITP EIRLLLTEF*
```

m135/a135 98.4% identity in 369 aa overlap

	10	20	30	40	50	60
m135.pep	MKYKRIVFKVGTSSITHSDGSLRSGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG					
a135	MKYKRIVFKVGTSSITHSDGSLRSGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG					
	10	20	30	40	50	60
m135.pep	FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL					
a135	FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL					
	70	80	90	100	110	120
m135.pep	SVLLQRRRAVP IINENDTVSVEELKIGDNDT LSAQVAAMIQADLLVLLTDIDGLYTGNPNS					
a135	SVLLQRRRAVP IINENDTVSVEELKIGDNDT LSAQVAAMIQADLLVLLTDIDGLYTGNPNS					
	130	140	150	160	170	180
m135.pep	NPDAVRLDKIEHINHEIEMAGGSGSANGTGGMLTKIKAATIAESGVVPYICSSLKPDA					
a135	NPDAVRLDKIEHINHEIEMAGGSGSANGTGGMLTKIKAATIAESGVVPYICSSLKPDA					
	190	200	210	220	230	240

385

	250	260	270	280	290	300
m135.pep	LAEEAAEHQADGSFFVPRAGLRTQKQWLAFYSESRGVSIVDEGAEHALSEQKSLMSGI					
a135	LAEEAADNQADGSFFVPRAGLRTQKQWLAFYSESRGGVYVDEGAEHALSEQKSLMSGI					
	250	260	270	280	290	300
	310	320	330	340	350	360
m135.pep	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAAEDLLKSRKAKGVFIHRDDWISITP					
a135	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAAEDLLKLRKAKGVFIHRDDWISITP					
	310	320	330	340	350	360
	370					
m135.pep	EIRLLLTEFX					
a135	EIRLLLTEFX					
	370					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 523>:

```

g136.seq
1  ATGGAAATCC GGTTCAGAC AGCATTTTTC CGTTTGGTTC AGatgaAAAC
51  AAACGCTtca aTTCTtaccg caACACGCCT TGTATTTCTT GccgCTGCCG
101 CACGGACAGG GATCGTTCCT GCCGgtTTTT TCCCCTTCCC TGCGGACGGT
151 TTGCGGTTTG TTGATGACCG CCTGCCAGTA GCGGTAGATG Tctgccagcg
201 cgTAAGGCag tTCGGACgca agttccgcca gctcgccttc ggTGAATTGC
251 AGgcgataa cgccgtttTC CTCTTCGTCg taaatgcgcg ccactgccat
301 cacgGGGTAA AACAGCTCTT CAAACGCTTC ATCATCGGCG GCTTCAAACC
351 AATCGGTCCG CACAATGTCC AAACCGTAAA GATAGGCGTT GCACCAAGTG
401 TAAAAATCGC TGCCGCCCTC GCCGTCTGTC TAGAGCCACA AATCGGGCAG
451 CTTTTTATCC GACATCGCGG CGGTTGTTTC CATCGCCATT GCCAAAACCA
501 GCCGTTCGAT TTCGGAACGT TCGGCGCGCG TAAATTGCGA TTCGTGCCCC
551 AACACTTCGG GCAGCCAGTC GAGCGGTGCC AATTGTCCG GCCCGCTCAA
601 CAGCGCCGTC ATAAACCTT GAACCTCGTC GCAACGCATC GTGTGCGCTT
651 GTTCGCTTTT GGCATCCAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 524; ORF 136.ng>:

```

g136.pep
1  MEIRFQTAFL RLVQMKTNAS ILTATRLVFP AAAARTGIVP AGFFPFPADG
51  LRFVDDRLPV AVDVCQRVRQ FGRKFRQLAF GELQADNAVF LFFVNAAHCH
101 HGVKQLFKRF IIGGFKPIGR HNVQTVKIGV APSVKIAAAL AVVVEPQIGQ
151 LFIRHRGGCF HRHCQNQPFQ FGTFGGGLR FVAQHFGQPV ERCQFVRPAQ
201 QRRHKTNLV ATHRVALFAF GIQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 525>:

```

m136.seq
1  ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
51  CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG
101 CGGACGGTTT GCGGTTTGTG GATGACTGCC TGCCAGTAGC GGTAGATATC
151 CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG
201 TGAATTGCAG ACGGATAGCG CCGTTTTTCT CTTTCGTCGT AATACCGCCC
251 AATGCCATGA TGGGATAAAA CAACTCTTCA AACGTTTCAT CATCGACGGC
301 TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC
351 ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTTATA CAGCCACAAA
401 TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTCCTA TCGCCATTGC
451 CAAAACAGC CGTTCGATTT CGGAACGTTT GGCGGCGGTA AATTGCGATT
501 CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC
551 CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
601 GTTGCCTTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG
651 ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCCTTTT CAAATGGGTT
701 TTGCGCCCTA TTATCGCCGC AATGCCGTCT GA

```

This corresponds to the amino acid sequence <SEQ ID 526; ORF 136>:

386

m136.pep
 1 METNASILTA TRLVFSAAAA RTGIVPACFF AFPADGLRFV DDCLPVAVDI
 51 RQCIRQLGFQ FRQLAFCELO TDSAVFLFVV NTAQCHDGIK QLFKRFIIDG
 101 FKPIGRHNIQ TVKISIAPCV KIAAAVFVFI QPQIGQFFIR HRGGCFHRHC
 151 QNQPFDFGTF GGGKLRFVAQ HFGQPVERCQ FVRPAQQRH KTLNVLVATHR
 201 VALFAFGIQQ FAQPPFGCFG KFSGIHHFPF QMGFAPYYRR NAV*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 136 shows 85.6% identity over a 209 aa overlap with a predicted ORF (ORF 136.ng) from *N. gonorrhoeae*:

m136/g136

		10	20	30	40
m136.pep		METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRFVDDCLPV			
		:			
g136	MEIRFQTAFLRLVQMKTNASILTATRLVFPAARTGIVPAGFFFPADGLRFVDDRLPV	10	20	30	40
		50	60	70	80
m136.pep	AVDIRQCIRQLGFQFRQLAFCELO TDSAVFLFVVNTAQCHDGIKQLFKRFIIDGFKPIGR				
	: : : : : :				
g136	AVDVCQVRVQFGRKFRQLAFGELQADNAVFLFVNAAHCHGVKQLFKRFIIGGFKPIGR	70	80	90	100
		110	120	130	140
m136.pep	HNIQTVKISIAPCVKIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLR				
	: : : :				
g136	HNVTQVKIGVAPSVKIAAALAVVVEPQIGQLFIRHRGGCFHRHCQNQPFDFGTFGGGKLR	130	140	150	160
		170	180	190	200
m136.pep	FVAQHFGQPVERCQFVRPAQQRHKTNLVLVATHRVALFAFGIQQFAQPPFGCFGKFSGIH				
g136	FVAQHFGQPVERCQFVRPAQQRHKTNLVLVATHRVALFAFGIQX	190	200	210	220
		230	240		
m136.pep	HFPFQMGFAPYYRRNAVX				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 527>:

a136.seq
 1 ATGGAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
 51 CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG
 101 CGGACGGTTT GCGGCTTGTT GATGACCGCC TGCCAGTAGC GGTAGATATC
 151 CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG
 201 TGAATTGCAG ACGGATAGTG CCGTTGTCCT CTTCGTCGTA AATACCGCCC
 251 AATGCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC
 301 TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC
 351 ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTTATA CAGCCACAAA
 401 TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCCATTGC
 451 CAAAACAGC CGTTCGATT CGGAACGTT GCGGCGGTA AATTGCGATT
 501 CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC
 551 CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
 601 GTTGCTTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG
 651 ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCTTTT CCAATGGGTT
 701 TTGCGCCCTA TTATAGTGA TTAAATTTAA ATCAGGACAA GGCGACGAAG
 751 CCGCAGACAG TACAAATAGT ACGGCAAGGC GAGGCAACGC CGTACTGGTT
 801 TAAATTTAAT CCACTATATC GCCGCAATGC CGTCTGA

This corresponds to the amino acid sequence <SEQ ID 528; ORF 136.a>:

a136.pep

387

```

1  METNASILTA  TRLVFSAAAA  RTGIVPACFF  AFPADGLRLV  DDRLPVAVDI
51  RQCIRQLGFQ  FRQLAFCELQ  TDSAVVLFVV  NTAQCHDGIK  QLFKRFIIDG
101 FKPIGRHNIQ  TVKISIAPCV  KIAAAVFVFI  QPQIGQFFIR  HRGGCFHRHC
151 QNQPFDFGTF  GGGKLRFVAQ  HFGQPVERCQ  FVRPAQQRH  KTLNLVATHR
201 VALFAFGIQQ  FAQPPFGCFG  KFSGIHHFFP  PMGFAPYYSG  LNLNQDKATK
251 PQTVQIVRQG  EATPYWFKFN  PLYRRNAV*

```

m136/a136 98.3% identity in 238 aa overlap

	10	20	30	40	50	60
m136.pep	METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRFVDDCLPVAVDIRQCIRQLGFQ					
a136	METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRLVDDRPLPVAVDIRQCIRQLGFQ					
	10	20	30	40	50	60
	70	80	90	100	110	120
m136.pep	FRQLAFCELQTD SAVVLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV					
a136	FRQLAFCELQTD SAVVLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m136.pep	KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ					
a136	KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ					
	130	140	150	160	170	180
	190	200	210	220	230	240
m136.pep	FVRPAQQRHKTNLNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFFPQMGFAPYYRR					
a136	FVRPAQQRHKTNLNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFFPQMGFAPYYSG					
	190	200	210	220	230	240
m136.pep	NAVX					
a136	LNLNQDKATKPQTVQIVRQGEATPYWFKFNPLYRRNAVX					
	250	260	270			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 529>:

```

g137.seq
1  ATGATTATCC  ATCACcaATT  CGATCCCGTC  CTCATCAGTA  TCGGCCCGCT
51  TGCCGTCCGC  TGGTATGCCT  TAAGCTACAT  CCTCGGATTT  ATTCTTTTTA
101 CCTTTCTCGG  CAGAAGCGCG  ATCGCGCAAG  GCTTGTCCTG  TTTTACCAAA
151 GAATCGCTCG  ACGACTTCCT  GACATGGGGC  ATTTTGGGCG  TGATTTTGGG
201 CGGACGCTTG  GGCTATGTCC  TGTTTTACAA  ATTCTCCGAC  TACCTCGCCC
251 ATCCGCTTGA  TATTTTCAAG  GTATGGGAAG  GCGGAATGTC  GTTCCACGGC
301 GGCTTTTGG  GTGTAGTTAT  TGCCATATGG  TTGTTAGCC  GCAAGCACGG
351 CATCGGCTTC  CTCAAAGTGA  TGGACACGGT  CGCGCCGCTC  GTTCCGCTGG
401 GTCTCGCTTC  GGGACGTATC  GGCAACTTTA  TCAACGGCGA  ACTTTGGGGA
451 CGCATTACCG  ACATTAAACG  ATTTTGGGCA  ATGGGCTTCC  CGCAAGCGCA
501 TTACGAAGAT  GCCGAAGCCG  CCGCGCACAA  TCCGCTTGG  GCAGAATGGC
551 TGCAACAATA  CGGTATGCTG  CCGCGTCATC  CCTCGCAGCT  TTATCAGTTT
601 GCCCTTGAAG  GCATCTGCCT  GTTCGCCGTC  GTTTGGCTGT  TTTCCAAAAA
651 ACCGCGCCCG  ACCGGGCGAG  CTGCCGCGCT  TTTTCTCGGC  GGCTACGGCG
701 TGTTCCGCTT  TATTGCCGAA  TTTGCGCGCC  AACCCGACGA  CTATCTCGGG
751 CTGCTGACCT  TGGGGCTGTC  GATGGGGCAA  TGGTTAGCG  TCCCGATGAT
801 TGTTTTGGGT  ATCGTCGGCT  TTGTCCGTT  CGGCATGAAA  AAACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 530; ORF 137.ng>:

g137.pep

m137.seq

m137.pep

m137/g137

	10	20	30	40	50	60
m137.pep	MITHPQFDPVLISIGPLAVRWYALS	SYILGFILFTFLGRRRIAQG	LSVFTKESLD	DFLT	TWG	
g137	MIIHQFDPVLISIGPLAVRWYALS	SYILGFILFTFLGRRRIAQG	LSVFTKESLD	DFLT	TWG	
	10	20	30	40	50	60
	70	80	90	100	110	120
m137.pep	ILGVILGGRLGYVLFYKFSDYL	AHPLDIFKVWEGGMSFHGG	FLGVVIAIRLFG	RKHGIG	F	
					:	
g137	ILGVILGGRLGYVLFYKFSDYL	AHPLDIFKVWEGGMSFHGG	FLGVVIAIWLFSR	KHGIG	F	
	70	80	90	100	110	120
	130	140	150	160	170	180
m137.pep	LKLMDTVAPLVPLGLASGRIGN	FINGELWGRVTDINAFWAMG	FPQARYEDAEAAAHNPLW			
			:	:		
g137	LKLMDTVAPLVPLGLASGRIGN	FINGELWGRITDINAFWAMG	FPQAHYEDAEAAAHNPLW			
	130	140	150	160	170	180
	190	200	210	220	230	240
m137.pep	AEWLQQYGMLPRHP	SQLYQFALEGICLFTVI	WLFSKKORSTGOVASL	FLGGYGIFRFIAE		

389

```

|||||
g137      AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSSKKPRPTGQTAALFLGGYGIVFRFIAE
          190      200      210      220      230      240
          250      260      270      280
m137.pep  FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFSGMKKQHX
          |||||
g137      FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFSGMKKQHX
          250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 533>:

```

a137.seq
1  ATGATTACCC ATCCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
51  TGCCGTCGCG TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTTAA
101 CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151 GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TAATTTTGGG
201 CGGGCGTTTG GGTACGTCC TGTTTTACAA GTTTTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTGGG GTGTAGTTAT TGCCATATGG TTGTTCCGTC GCAAACACGG
351 CATCGGCTTC CTCAAACCTGA TGGACACGGT CGCACCGCTC GTTCCACTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTCA TCAACGGCGA ACTTTGGGGA
451 CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG
501 TTACGAAGAC CTCGAAGCCG CCGCGCACAA TCCGCTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCACCTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TCTCTAAAAA
651 ACAGCGGCCG ACCGGACAAG TCGCCTCACT CTTCTCGGC GGCTACGGCA
701 TATTCCGCTT CATTGCCGAA TTTGCACGCC AACCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGTTGAGCG TCCCAGATGAT
801 TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 534; ORF 137.a>:

```

a137.pep
1  MITHPQFDPV LISIGPLAVR WYALSYILGF ILFTFLGRRR IAQGLSVFTK
51  ESLDDFLTGW ILGVILGGR L GYVLFYKFS YLAHPLDIFK VWEGGMSFHG
101 GFLGVVIAIW LFGKKGIGF LKLMDTVAPL VPLGLASGRI GNFINDELWG
151 RVTDINAFWA MGFPQARYED LEAAAHNPLW AEWLQQYGML PRHPSQLYQF
201 ALEGICLFAV VWLFSKKQRP TGQVASLFLG GYGIFRFIAE FARQPDDYL
251 LLTLGLSMGQ WLSVPMIVLG IVGFVRFSGM KQH*

```

m137/a137 98.2% identity in 283 aa overlap

```

          10      20      30      40      50      60
m137.pep  MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTGW
          |||||
a137      MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTGW
          10      20      30      40      50      60
          70      80      90      100     110     120
m137.pep  ILGVILGGR LGYVLFYKFS DYLAHPLDIFK VWEGGMSFHGGFLGVVIAIRLFGKKGIGF
          |||||
a137      ILGVILGGR LGYVLFYKFS DYLAHPLDIFK VWEGGMSFHGGFLGVVIAIRLFGKKGIGF
          70      80      90      100     110     120
          130     140     150     160     170     180
m137.pep  LKLMDTVAPLVPLGLASGRIGNFINDELWGRVTDINAFWAMGFPQARYEDAEAAAHNPLW
          |||||
a137      LKLMDTVAPLVPLGLASGRIGNFINDELWGRVTDINAFWAMGFPQARYEDAEAAAHNPLW
          130     140     150     160     170     180
          190     200     210     220     230     240
m137.pep  AEWLQQYGMLPRHPSQLYQFALEGICLFTVIWLFSSKKQRSTGQVASLFLGGYGIFRFIAE
          |||||
a137      AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSSKKQRPTGQVASLFLGGYGIFRFIAE

```

390

	190	200	210	220	230	240
	250	260	270	280		
m137.pep	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQH	X				
a137	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQH	X				
	250	260	270	280		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 535>:

g138.seq

```

1  ATGGAGTTTG AAAACATTAT TTCCGCCGCG gaCAAGGCGC GTATCCTTGC
51  CGAAGCACTG CCTTACATcc gccgGTTTTTC CGGTTCCGGTC GCCGTCATCA
101 AGTATGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGA AAAAGTC GGCAAAAAGG
251 GCGAATTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGAC GATGGATATT
301 GTCGAAATGG TATTGGCGG GCACGTCAAC AAGGAAATCG TGTCGATGAT
351 TAACACATAT GGAGGCGACG CGGTCGGCGT GAGCGGGCGC GACGACCATT
401 TCATTAAGGC GAAGAACTT TTGGTCGATA CGCCCGAACA GAATAGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGCGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT GGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAAtatcgc
651 cgGTGTGATG GACAAAACGG GCAATCTGCT GACCAAATC acgCCGAAAC
701 GGATTGATGG GCTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCcgccgtc aACGGTGTGA AAGCCACGCA
801 CATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGGTCGATG ATTTTAGGCA GAGGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 536; ORF 138.ng>:

g138.pep

```

1  MEFENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51  RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKEFEVQGM RVTDKETMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LVDTPQNSV
151 DIGQVGTVES IDTGLVKGLI ERGCI PVVAP VGVGEKG EAF NINADLVAGK
201 LAEELNAEKL LMMTNIA GVM DKTGNLLTKL TPKRIDGLIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGRGEDA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 537>:

m138.seq

```

1  ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
51  CGAAGCGCTG CCTTACATCC GCCGTTTTTC CGGTTCCGGTC GCCGTCATCA
101 AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGA AAAAGTC GGCAAAAAGG
251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
301 GTCGAAATGG TGTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
351 TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
401 TCATTAAGGC GAAGAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT AGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAATATCGC
651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAATC ACGCCGAAAC
701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCCGTC AACGGTGTGA AAGCCACGCA
801 TATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGTTCGATG ATTTTGGGCG GTGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 538; ORF 138>:

m138.pep

```

1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA

```

51 RDVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKAMDI
 101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
 151 DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
 201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
 251 KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGGGEDA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 138 shows 98.0% identity over a 298 aa overlap with a predicted ORF (ORF 138.ng)
 from *N. gonorrhoeae*:

m138/g138

	10	20	30	40	50	60
m138.pep	MESENIIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVLLKLVG					
g138	MEFENIIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY					
g138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKETMDIVEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
g138	GGHAVGVSGRDDHFIKAKKLLVDTPEQNSVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
g138	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDGLIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX					
g138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGRGEDAX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 539>:

a138.seq

1	ATGGAGTCTG	AAAACATTAT	TTCCGCCGCC	GACAAGGCGC	GTATCCTTGC
51	CGAAGCGCTG	CCTTACATCC	GCCGGTTTTC	CGGTTCGGTC	GCCGTCATCA
101	AATACGGCGG	CAACGCGATG	ACCGAACCTG	CCTTGAAAGA	AGGGTTTGCC
151	CGCGATGTCT	TGCTGCTGAA	GCTGGTCGGC	ATTCATCCCG	TCATCGTTCA
201	CGCGCGCGGG	CCGCAGATCA	ATGCGATGCT	TGAAAAAGTC	GGCAAAAAGG
251	GTGAGTTTGT	CCAAGGAATG	CGCGTTACCG	ACAAAGAGGC	GATGGATATT
301	GTGAAAATGG	TGTTGGGCGG	GCATGTCAAT	AAAGAAATCG	TGTCGATGAT
351	TAACACATAT	GGCGGACACG	CGGTCGGCGT	AAGCGGACGC	GACGACCATT
401	TCATTAAGGC	GAAGAACTT	TTGATCGATA	CGCCCGAACA	GAATGGCGTG
451	GACATCGGAC	AGGTCGGTAC	GGTGGAAGC	ATCGATACCG	GTTTGGTTAA
501	AGGGCTGATA	GAACGTGGCT	GCATTCCCGT	CGTCGCCCCC	GTGCGCGTAG
551	GTGAAAAAGG	CGAAGCGTTC	AACATCAACG	CCGATTGGT	AGCAGGCAAA
601	TTGGCGGAAG	AATTGAACGC	CGAAAACTC	TTGATGATGA	CGAATATCGC
651	CGGTGTGATG	GACAAAACGG	GCAATCTGCT	GACCAAATC	ACGCCGAAAC
701	GGATTGATGA	ACTGATTGCC	GACGGCACGC	TGTATGGCGG	TATGCTGCCG
751	AAAATCGCTT	CTGCGGTCGA	AGCCGCCGTC	AACGGCGTGA	AAGCCACGCA
801	TATCATCGAC	GGCAGGGTGC	CCAACGCGCT	TTTGCTGGAA	ATCTTTACCG
851	ATGCCGGTAT	CGGTTCGATG	ATTTTGGGCG	GTGGGGAAGA	TGCCTGA

This corresponds to the amino acid sequence <SEQ ID 540; ORF 138.a>:

```
a138.pep
  1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
 51  RDVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKAEMDI
101  VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
151  DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
201  LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
251  KIASAVEAAV NGVKATHIID GRVFNALLLE IFTDAGIGSM ILGGGEDA*
```

m138/a138 99.7% identity in 298 aa overlap

	10	20	30	40	50	60
m138.pep	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVLLKLVG					
a138	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKAEMDIVEMVLGGHVNKEIVSMINTY					
a138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKAEMDIVEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
a138	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGEAFNINADLVAGKLAEEELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
a138	VGVGEKGEAFNINADLVAGKLAEEELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRVFNALLLEIFTDAGIGSMILGGGEDAX					
a138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRVFNALLLEIFTDAGIGSMILGGGEDAX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 541>:

```
g139.seq
  1  ATGCGAACCA CCTCAACCTT CCCTACAAAA ACTTCAAAC CGGCTGCCAT
 51  GGC GTTAGCT GTTGCAACAA CACTTCTGCT CTGCTTAggc ggcggcggag
101  gcGGCACTTC TGCTCCCGAC TTTAATGCAG GCGGCACCGG TATCGGCAGC
151  AACAGCAGGG CAACGATAGC GGAATCAGCA GCAGTATCTT ACGCCGGTAT
201  AAAAAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
251  ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAAAGCCCC CCGAATCTGC
301  ATACCGGAGA CTTTCAAAC CCAAATGACC AATATTAAGA ATATGATCAA
351  CCTCAAACCT GCAATTGAAG CAGGCTATAC AGGACGCGGG GTAGAGGTAG
401  GTATCGTCGA TACAGGCGAA TCCGTCGGCA GCATATCCTT TCCCGAACTG
451  TATGGCAGAA AAGAACACGG CTATAACGAA AATTACAAAA ACAAAATTACA
501  AAAACTATAC GGCGTATATG CGGAAGGAAG CGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 542; ORF 138.ng>:

```
g139.pep
  1  MRTTSTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
 51  NSRATIAESA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKIKAPRIC
101  IPETFQTQMT NIKNMINLKP AIEAGYTGRG VEVGIVDTGE SVGSISFPEL
151  YGRKEHGYNE NYKNKLQKLY GVYAEGSA*
```

m139.seq

```

m139.seq
1  ATGCGAACGA  CCCCAACCTT  CCCTACAAAA  ACTTTCAAAC  GACTGCCAT
51  GGCGTTAGCT  GTTGCAACAA  CACTTTCTGC  CTGCTTAGGC  GGCGGCGGAG
101 GCGGCACTTC  TGCGCCCGAC  TTCAATGCAG  GCGGTACCGG  TATCGGCAGC
151 AACAGCAGTC  CAACAACAGC  GAAATCAGCA  GCAGTATCTT  AGCCCGGTAT
201 CAAGAACGAA  ATGTGCAAAG  ACAGAAGCAT  GCTCTGTGCC  GGTCCGGATG
251 ACGTTGCGGT  TACAGACAGG  GATGCCAAAA  TCAATGCCCC  CCCCCGAATC
301 TGCATACCGG  AGACTTTCCA  AACCCAAATG  ACGCATtACA  AGAATTGAT
351 CAACCTCAA  CCTGCAATTG  AAGCAGGCTA  TACAGACGCG  GGGGTAGAGG
401 TAGGTATCGT  CGACACAGGC  GAATCCGTCG  GCAGCATATC  CTTTCCCGAA
451 CTGTATGGCA  GAAAGAACA  CGGCTATAAC  GAAAATTACG  AAAAATAATA
501 CGGCGTATAT  GCGGAAGGAA  GCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:

m139.pep

1	<u>MRTTPTFPTK</u>	<u>TFKPTAMALA</u>	VATTLSACLG	GGGGGTSAPD	FNAGGTGIGS
51	NSRATTAKSA	AVSYAGIKNE	MCKDRSMLCA	GRDDVAVTDR	DAKINAPPRI
101	CIPETFQTQM	THYKNLNLK	PAIEAGYTGR	GVEGVIVDTH	ESVSGISFPE
151	LYGRKEHGYN	ENYEKLYGVY	AEGSA*		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae*:

```

m139/g139

                10          20          30          40          50          60
m139.pep  MRTTPTFPKTKFKPTAMALAVATTLSACLGGGGGTSAPDFNAGGTGIGSNSRATTAKSA
           ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| : |||
g139      MRTTSTFPKTKFKPAAMALAVATTLSACLGGGGGTSAPDFNAGGTGIGSNSRATIAESA
           10          20          30          40          50          60

                70          80          90          100         110         120
m139.pep  AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPRICIPETFQTQMTHYKNLNLK
           ||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| : ||||| : |||||
g139      AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKIKAP-RICIPETFQTQMTNIKNMINLK
           70          80          90          100         110

                130         140         150         160         170
m139.pep  PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENY----EKLYGVYAEGSAX
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| |||||
g139      PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENYKNKLQKLYGVYAEGSAX
           120         130         140         150         160         170

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 545>:

```

a139.seq
1  ATGCGAACGA  CCCCAACCTT  CCCTACAAAA  ACTTTCAAAC  CGGCTGCCAT
51  GGCCTTAGCT  GTTGCAACAA  CACTTCTCTC  CTGCTTAGGC  GGCGCGGGAG
101 GCGGCACTTC  TGCGCCCGAC  TTCAATGCAG  GCGGCACCGG  TATCGGCAGC
151 AACAGCAGGG  CAACAACGAC  GAAATCAGCA  GCAATATCTT  AGCCCGGTAT
201 CAGAAGCAG  GTTGCAAAAG  ACAGAAGCAT  GCTCTGTGCC  GGTCGGGATG
251 ACGTTGCGGT  TACAGACAGG  GATGCCAAAA  TCAATGCCCC  CCCCGAATC
301 TGCATACCGG  AGACTTTACA  AACCCAAATG  ACGCAT.ACA  AGAATTTGAT
351 CAACCTCAA  CCTGCAATTG  AAGCAGGCTA  TACAGGACGC  GGGGTAGAGG
401 TAGGTATCGT  CGACACAGGC  GAATCCGTCG  GCAGCATATC  CTTTCCCGAA
451 CTGTATGGCA  GAAAAGAAAC  CGGCTATAAC  GAAAATTAC.  AAAAATAATA
501 CGGCGTATAT  GCGGAAGGAA  GCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>:

a139.pep
1 MRTTPTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS

394

51 NSRATTAKSA AISYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPRI
 101 CIPETLQTQM THXKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSI SFPE
 151 LYGRKEHGYN ENYXKLYGVY AEGSA*

m139/a139 97.1% identity in 175 aa overlap

	10	20	30	40	50	60
m139.pep	MRTTPTFPTKTFKPTAMALAVATTL	SACLG	GGGGG	TSAPDFNAGGTGIGSNSRATTAKSA		
a139	MRTTPTFPTKTFKPAAMALAVATTL	SACLG	GGGGG	TSAPDFNAGGTGIGSNSRATTAKSA		
	10	20	30	40	50	60
	70	80	90	100	110	120
m139.pep	AVSYAGIKNEMCKDRSMLCAGRDDVAVTDR	DAKINAPPRICIPET	FQTMTHXKNLINLK			
a139	AISYAGIKNEMCKDRSMLCAGRDDVAVTDR	DAKINAPPRICIPET	LQTQMTHXKNLINLK			
	70	80	90	100	110	120
	130	140	150	160	170	
m139.pep	PAIEAGYTGRGVEVGIVDTGESVGSISFPE	LYGRKEHGYNENYKLYGVY	AEGSAX			
a139	PAIEAGYTGRGVEVGIVDTGESVGSISFPE	LYGRKEHGYNENYXKLYGVY	AEGSAX			
	130	140	150	160	170	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 547>:

g140.seq

1	Atgtcggcac	gCGGCAAGGG	GGCAGgetat	ctcAACAGTA	CCGGACGACa
51	TGTTCCCTTC	CTGAGTGCCG	CCAAAATCGG	GCAGGATTAT	TCTTTCTTCA
101	AAAATATCAA	AACCGACGGC	GGTCTGCTGG	CTTCCCTCGA	CAGCGTCGAA
151	AAAACAGCGG	GCAGTGAAGG	CGACACGCCG	TCCTATTATG	TCCGTCGCGG
201	CAATGCGGCA	CGACTGCTT	CGGCAGCGGC	ACATTCCGCG	CCGCGCGGTC
251	TGAAACACGC	CGTAGAACAG	GGCGGCAGCA	ATCTGGAAAA	CCTGATGGTC
301	GAGCTGGATG	CCTCCGAATC	ATCCGCAACA	CCCGAGACGG	TTGAAACTGC
351	GGTCGCCGAC	CGCACAGATA	TGCCGGGCAT	CCGCCTACGG	CGCACAACTT
401	TCCGCACAGC	GGCAGCCGTA	CAGCATGCGA	ATACCGCCGA	CGGCGTACGc
451	aTCTTcaaCA	GTCTCGCCGC	TAccgTCTAt	GccgACAGTG	CCGCCGCCCA
501	TGccgATATG	CAGGGACGCC	GCCTGAAAGC	CGTATCGGAC	GGGTGGACC
551	ACAACGGTAC	GGGTCTGCGC	GTCAATCGCGC	AAACCCAACA	GGACGGTGGA
601	ACGTGGGAAC	AGGGCGGTGT	CGAAGGCAAA	ATGCGCGGCA	GTACCCAAAC
651	TATCGGCATT	GCCGCGAAAA	CCGGCGAAAA	TACGACAGCA	GCCGCCACAC
701	TGGGCATAGG	ACGCAGCACA	TGGAGCGAAA	ACAGTGCAAA	TGCAAAAACC
751	GACAGCATTa	GTCTGTTTGC	AGGCATACGG	CACGATGTGG	GCGATATCGG
801	CTATCTCAAA	GGCCTGTTCT	CctaCGGACG	CTACAAAAAC	AGCATCAGCC
851	GCAGCACCAG	TGCGGATGAA	TATGCGGAAG	GCAGCGTCAA	CGGCACGCTG
901	ATGCAGCTGG	GCGCACTGGG	TGGTGTCAAC	GTTCCGTTTG	CCGCAACGGG
951	AGATTTGACG	GTTGAAGGCG	GTCTGCGCCA	CGACCTGCTC	AAACAGGATG
1001	CATTTCGCCGA	AAAAGGCagt	GCTTTGGGCT	GGAGCGGCAA	CAGCCTCACT
1051	GAAGGCACAC	TGGTCGGACT	CGCGGTCTG	AAACTGTGCG	AACCCCTGAG
1101	CGATAAAGCC	GTCCTGTCTG	CGACGGCGGG	CGTGGAACGC	GACCTGAACG
1151	GACGCGACTA	CGCGGTAACG	GGCGGCTTTA	CCGCGCGGGC	TGCAGCAACC
1201	GGCAAGACGG	GTGCACGCAA	TATGCCGCAC	ACCCGCCGGG	TTGCCGGTCT
1251	GGGGGTGGAT	GTCGAATTCT	GCAACGGCTG	GAACGGCTTG	GCACGTTACA
1301	GCTACACCGG	TTCCAAACAG	TACGGCAACC	ACAGCGGACA	AATCGGCGTA
1351	GGCTACCGGT	TCTGA			

This corresponds to the amino acid sequence <SEQ ID 548; ORF 140.ng>:

g140.pep

1	MSARGKGAGY	LNSTGRHVPF	LSAAKIGQDY	SFFKNIKTDG	GLLASLDSVE
51	KTAGSEGDTP	SYVVRGNAA	RTASAAHSA	PAGLKHAVEQ	GGSNLENLMV
101	ELDAESSAT	PETVETAVAD	RTDMPGIRLR	RTTFRTAAAV	QHANTADGVR
151	IFNSLAATVY	ADSAAAHADM	QGRRLKAVSD	GLDHNGTGRL	VIAQTQDGG
201	TWEQGGVEGK	MRGSTQTIGI	AAKTGENTTA	AATLGIGRST	WSENSANAKT

251 DSISLFAGIR HDVGDIGYLK GLFSYGRYKN SISRSTGADE YAEGSVNGTL
 301 MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT
 351 EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAAT
 401 GKTGARNMPH TRRVAGLGVD VEFNGWNGL ARYSYTGSKQ YGNHSGQIGV
 451 GYRF*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 549>:

m140.seq
 1 ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG
 51 TGTTCCTTC CTGAGTGCCG CCAAATCGG GCAGGATTAT TCTTCTTCA
 101 CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCTCGA CAGCGTCGAA
 151 AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG
 201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC
 251 TGAAACACGC CGTAGAACAG GGCGGCAGCA ATCTGGAAAA CCTGATGGTC
 301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
 351 GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCTAC GGCGCAACTT
 401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CCGGTGTACGC
 451 ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA
 501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
 551 ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCAACA GGACGGTGGA
 601 ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCAAAAC
 651 CGTCGGCATT GCCGCGAAAA CCGGCGAAAA TACGACAGCA GCCCGCACAC
 701 TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
 751 GACAGCATTG GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG
 801 CTATCTCAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC
 851 GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CCGCACGCTG
 901 ATGCAGCTGG GCGCACTGGG CCGTGTCAAC GTTCCGTTTG CCGCAACGGG
 951 AGATTTGACG GTCGAAGCGG GTCTGCGCTA CGACCTGCTC AAACAGGATG
 1001 CATTGCGCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCCTCACT
 1051 GAAGGCACGC TGGTCGGACT CGCGGGTCTG AAGCTGTGCG AACCTTGAG
 1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
 1151 GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
 1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT
 1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
 1301 GATACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
 1351 GGCTACCGGT TCTGA

This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:

m140.pep
 1 MSARGKGAGY LNSTGRRVPF LSAKIGQDY SFFTNIETDG GLLASLDSVE
 51 KTAGSEDTL SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
 101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR
 151 IFNSLAATVY ADSTAAHADM QGRRLLKAVSD GLDHNGTGLR VIAQTQDDGG
 201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT
 251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
 301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT
 351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
 401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV
 451 GYRF*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 140 shows 94.5% identity over a 454 aa overlap with a predicted ORF (ORF 140.ng) from *N. gonorrhoeae*:

m140/g140

	10	20	30	40	50	60
m140.pep	MSARGKGAGYLNSTGRRVPFLSAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEDTL					
	: : : : :					
g140	MSARGKGAGYLNSTGRHVPFLSAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEDTLP					
	10	20	30	40	50	60
	70	80	90	100	110	120

396

m140.pep	SYVVRGNAARTASAAAHSA	PAGLKHAVEQGGSNLENLM	VELDASESSATPETVETAAAD
g140	SYVVRGNAARTASAAAHSA	PAGLKHAVEQGGSNLENLM	VELDASESSATPETVETAVAD
	70	80	90 100 110 120
m140.pep	RTDMPGIRPYGATFRAAAAV	QHANAADGVRIFNSLAATVYAD	STAAHADMQGRRRLKAVSD
g140	RTDMPGIRLRRTTFR	TAAAVQHANTADGVRIFNSLAATVYAD	SAAAHADMQGRRRLKAVSD
	130	140	150 160 170 180
m140.pep	GLDHNGTGLRVIAQTQDGGT	WEQGGVEGKMRGSTQT	VGIAAKTGENTTAAATLGMGRST
g140	GLDHNGTGLRVIAQTQDGGT	WEQGGVEGKMRGSTQT	IGIAAKTGENTTAAATLGIGRST
	190	200	210 220 230 240
m140.pep	WSENSANAKTDSISLFAGIR	HDAIGYLGKLFSGRYKNSISR	STGADEHAEGSVNGTL
g140	WSENSANAKTDSISLFAGIR	HVDIGYLGKLFSGRYKNSISR	STGADEYAEGSVNGTL
	250	260	270 280 290 300
m140.pep	MQLGALGGVNVFPAATGDLT	VEGGLRYDLLKQDAFAEKGSAL	GWSGNSLTEGTLVGLAGL
g140	MQLGALGGVNVFPAATGDLT	VEGGLRHDLLKQDAFAEKGSAL	GWSGNSLTEGTLVGLAGL
	310	320	330 340 350 360
m140.pep	KLSQPLSDKAVLFATAGVER	DLNGRDYTVTGGFTGATAAT	GKTGARNMPHTRLVAGLGAD
g140	KLSQPLSDKAVLSATAGVER	DLNGRDYAVTGGFTGAAAAT	GKTGARNMPHTRRVAGLGVD
	370	380	390 400 410 420
m140.pep	VEFGNGWNLARYSYAGSKQY	GNHSGRVGVGYRFX	
g140	VEFGNGWNLARYSYTGSKQY	GNHSGQIGVGYRFX	
	430	440	450

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 551>:

a140.seq	1	ATGTCGGCAG	GCGGTAAGGG	GGCAGGCTAT	CTCAACCGTA	CCGGACAACG
	51	TGTTCCCTTC	CTGAGTGCCG	CCAAAATCGG	GCGGGATTAT	TCTTTCTTCA
	101	CAAACATCGA	AACCGACGGC	GGTCTGCTGG	CTTCCCTCGA	CAGCGTCGAA
	151	AAAACAGCGG	GTAAGTGAAG	CGACACGCTG	TCCTATTATG	TCCGTCGCGG
	201	CAATGCGGCA	CGGACTGCTT	CGGCAGCGGC	ACATTCCGCG	CCCGCCGGTC
	251	TGAAACACGC	CGTAGAACAG	GGCGGCAGCA	ATCTGGAAAA	CCTGATGGTC
	301	GAACTGGATG	CCTCCGAATC	ATCCGCAACA	CCCGAGACGG	TTGAAACTGC
	351	GGCCGCCGAC	CGCACAGATA	TGCCGGGCAT	CCGCCCTAC	GGCGCAACTT
	401	TCCGCGCAGC	GGCAGCCGTA	CAGCATGCGA	ATGCCGCCGA	CGGTGTACGC
	451	ATCTTCAACA	ATCTCGCCGC	TACCGTCTAT	GCCGACAGTA	CCGCCGCCCA
	501	TGCCGATATG	CAGGGACGCC	GCCTGAAAGC	CGTATCGGAC	GGGTGTGGAC
	551	ACAACGCTAC	GGGTCTGCGC	GTCATCGCGC	AAACCCAACA	GGACGGTGGA
	601	ACGTGGGAAC	AGGGCGGTGT	TGAAGGCAAA	ATGCGCGGCA	GTACCCAAAC
	651	CGTCGGCATT	GCCGCGAAAA	CCGCGCAAAA	TACGACAGCA	GCCGCCACAC
	701	TGGGCATGGG	ACACAGCACA	TGGAGCGAAA	ACAGTGCAAA	TGCAAAAACC
	751	GACAGCATT	GTCTGTTTGC	AGGCATACGG	CACGATGCGG	GCGATATCGG
	801	CTATCTCAA	GGCCTGTTCT	CCTACGGACG	CTACAAAAAC	AGCATCAGCC
	851	GCAGCACC	GGTGCGGAA	CATGCGGAAG	GCAGCGTCAA	CGGCACGCTG
	901	ATGCAGCTGG	GCGCACTGGG	CGGTGTCAAC	GTTCCGTTTG	CCGCAACGGG

```

951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AACAGGATG
1001 CATTTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCATCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAGCTGTGCG AACCCCTGAG
1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGAACGC GACCTGAACG
1151 GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGCCTGG TTGCCGGTCT
1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301 GCTACGCCGG TTCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351 GGCTACCGGT TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 552; ORF 140.a>:

```

a140.pep
  1 MSAGGKGAGY LNRTGQRVFP LSAAKIGRDY SFFTNIETDG GLLASLDSVE
 51 KTAGSEGDTL SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR
151 IFNNLAATVY ADSTAAHADM QGRRRLKAVSD GLDHNATGLR VIAQTQDGG
201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGHST WSENSANAKT
251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSIT
351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV
451 GYRF*

```

m140/a140 98.2% identity in 454 aa overlap

m140.pep	10	20	30	40	50	60
	MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL					
a140	MSAGGKGAGYLNRTGQRVFPFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL					
	10	20	30	40	50	60
m140.pep	70	80	90	100	110	120
	SYVVRGNAAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
a140	SYVVRGNAAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
	70	80	90	100	110	120
m140.pep	130	140	150	160	170	180
	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNLAATVYADSTAAHADMQGRRRLKAVSD					
a140	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNLAATVYADSTAAHADMQGRRRLKAVSD					
	130	140	150	160	170	180
m140.pep	190	200	210	220	230	240
	GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGRST					
a140	GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHST					
	190	200	210	220	230	240
m140.pep	250	260	270	280	290	300
	WSENSANAKTDSISLFAGIRHDAGDIGYLGKGLFSYGRYKNSISRSTGADEHAEGSVNGTL					
a140	WSENSANAKTDSISLFAGIRHDAGDIGYLGKGLFSYGRYKNSISRSTGADEHAEGSVNGTL					
	250	260	270	280	290	300
m140.pep	310	320	330	340	350	360
	MQLGALGGVNPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLITEGLVGLAGL					
a140	MQLGALGGVNPFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITEGLVGLAGL					
	310	320	330	340	350	360
m140.pep	370	380	390	400	410	420
	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					
a140	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					

	370	380	390	400	410	420
	430	440	450			
m140.pep	VEFGNGWNLARYSYAGSKQYGNHSGRVGVGYRFX					
a140	VEFGNGWNLARYSYAGSKQYGNHSGRVGVGYRFX					
	430	440	450			

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 553>:

g141.seq

```

1   atgagcttca aAaccgATGC CGAAACCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAGCTG
151 CCGCAAAAAC AAGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TGC CGGAGCC TTCTTTGGGT
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GCGGCGGCT AC GCGCAAGT
351 AGAGCGTTtt gGCAATATTC ACCTGCACTT CACCGGCGAC TTCCACGCCA
401 TCGGTGCGGC GAATAACCTC CTCGCCGCCA TGCTCGACAA CCATATCTAC
451 CAAGGTAACG AGTTGAACAT CGACCCCAA CGCGTGCTGT GCGGCGCGCT
501 GGTGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGTATGGGCA
551 AGCCTGTtga cggCGTGATG CGtcccGACG GCTTCGACAT CACCGTCGCC
601 TCCGAAGTGa tggcgGTATT CTGCCTTGCC AAAGACATCA GCGATTGAA
651 AGAGCGTTttt gGCAATATTC TCGTCGCCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTGAAG GCACACGGCG CGATGGCGGC ATTGCTAAAA
751 GATGCGATTA AGCCCAATTT GGTGCAAACC ATCGAAGGCA CTCCGGCCTT
801 TGTACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTTA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACC GAAGCA
901 GGCTTCGGCG CGGACTTGGG TGCGGAAAAA TTCTGCGACA TCAATGCCG
951 CTTGCCCGGT TTGAAACCTG ATGCGGCACT CGTCGTGGCG ACTGTCCGCG
1001 CCCTGAAATA CAACGGCGGC GTGGAACGCG CCAACCTTGG TGAAGAAAC
1051 CTCGAAGCCT TGGCAAAAGG TTTGCCCAAC CTGTTGAAAC ACATTCCAA
1101 CCTGAAAAAC GTATTTCGAC TGCCCGTCTG CGTTGCGCTC AACC GCTTCG
1151 TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GCGGCGCGGG
1251 CCGCGCGGAT TTGGCGCGCA AAGTCGTCAA TGCCATCGAC AACC AACCTA
1301 ATAACCTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCCAAAAGT GTACGGCGCG GAAGATGTCG ATTT CAGCGC
1401 GGAAGCGTCT GCCGAAATCG CCTCGCTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATATT CATTGAGCGA CAACGCCAAA
1501 CTCTTGGGCT GCCCGAAGG CTTCCGCATC GCCGTACGCG GTATCACTGT
1551 TTCCGCGCGC GCGGGCTTCA TCGTTGCGTT GTGCGGCAAT ATGATGAAAA
1601 TGCCGGGCCT GCCGAAAGT CCGGCTGCCG AGAAAATCGA TGTGGACGAA
1651 CACGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 554; ORF 141.ng>:

g141.pep

```

1   MSFKTDAETA QSSTMRPIGE IAAKGLNVD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFVGKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPV DGVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERF GNILVAYAKD GSPVYAKDLK AHGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LEALAKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAID NQPNNF GFAY DVELGIKDKI
451 RAI AQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCP EGFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDE
551 HGVIHGLF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 555>:

m141.seq

```

1  ATGAGCTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTTGGTCT GAATGCCGAC AACATTGAGC
101 CTTACGGTCA TTACAAGGCG AAAATCAATC CTGCCGAAGC GTTCAAACCTG
151 CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GCGGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCGTTGC
251 GCCACATCGG CAAAGATGCC GTGATTGCCG TGCGCGAACC TTCTCTGGGG
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GCGGCGGCT ATGCCCAAGT
351 TTGCGCATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
401 TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAA CGCGTGCTGT GCGGCGCGT
501 GGTGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGTA
551 AACCCGTTGA CCGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTGAA
651 AGAGCGTTTG GGCAACATCC TTGTGCGCTA CGCCAAAGAC GGCAGCCCCG
701 TTACGCCCAA AGATTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
801 CGTACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CTTGCGCGT TGAAAACCTG ATGCGGCTGT TGTCGTGGCG ACTGTCCGCG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAAT
1051 TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCTAA
1101 CTGAAAAAAC GTATTGCGAC TGCCCGTCTG CGTTGCGCTC AACCGCTTCG
1151 TGTCGACGCG CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACTTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCCAAAAAGT GTACGGCGCG GAAGATGTTG ATTTACGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCCT
1551 TCCGCGAGGC GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651 GAAGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 556; ORF 141>:

m141.pep

```

1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNAD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDA VIALREPSLG
101 PVFGVKGGA GGGYAQVLP EDINLHFTGD FHAIGAANNL LAAMLNDNIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPVGDVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDADAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNGFAY DVELGIKDKI
451 RAIAQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 EGVIHGLF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 141 shows 97.5% identity over a 558 aa overlap with a predicted ORF (ORF 141.ng) from *N. gonorrhoeae*:

m141/g141

	10	20	30	40	50	60
m141.pep	MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL PQKQGRLLIV					
	:					
g141	MSFKTDAETAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKL PQKQGRLLIV					
	10	20	30	40	50	60

400

	70	80	90	100	110	120
m141.pep	TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGGAAGGGYAQVLPM					
g141	TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGGAAGGGYAQVLPM					
	70	80	90	100	110	120
	130	140	150	160	170	180
m141.pep	EDINLHFTGDFHAIGAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDMMNDRQLRNIID					
g141	EDINLHFTGDFHAIGAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDMMNDRQLRNIID					
	130	140	150	160	170	180
	190	200	210	220	230	240
m141.pep	GMGKPVGDGVMRPGFDITVASEVMAVFCLAKDISDLKERLGNILVAYAKDGSPVYAKDLK					
g141	GMGKPVGDGVMRPGFDITVASEVMAVFCLAKDISDLKERFGNILVAYAKDGSPVYAKDLK					
	190	200	210	220	230	240
	250	260	270	280	290	300
m141.pep	ANGAMAALLKDAIKPNLVQTIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
g141	AHGAMAALLKDAIKPNLVQTIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
	250	260	270	280	290	300
	310	320	330	340	350	360
m141.pep	GFGADLGAEKFCDIKCRLAGLKPDAAVVATVRALKYNGGVERANLGEENLDALEKGLPN					
g141	GFGADLGAEKFCDIKCRLAGLKPDAAVVATVRALKYNGGVERANLGEENLEALAKGLPN					
	310	320	330	340	350	360
	370	380	390	400	410	420
m141.pep	LLKHISNLKNVFGPLPVVVALNRFVSDADAELAMIEKACAEHGVEVSLTEVWGKGGAGGAD					
g141	LLKHISNLKNVFGPLPVVVALNRFVSDSDAELAMIEKACAEHGVEVSLTEVWGKGGAGGAD					
	370	380	390	400	410	420
	430	440	450	460	470	480
m141.pep	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
g141	LARKVVNAIDNQPNNGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
	430	440	450	460	470	480
	490	500	510	520	530	540
m141.pep	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
g141	LDKMPICMAKTQYSLSDNAKLLGCPGFRIVRGITVSAGAGFIVALCGNMMKMPGLPKV					
	490	500	510	520	530	540
	550	559				
m141.pep	PAAEKIDVDAEGVIHGLFX					
g141	PAAEKIDVDEHGVHGLFX					
	550					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 557>:

```

a141.seq
1  ATGAGTTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAACCTG
151 CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC

```

```

201 GCGGGGCGAA GGTAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TGCGCGAGCC TTCTTTGGGT
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GCGGCGGCT ATGCCCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTACGCCCA
401 TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAA CGCGTGCTGT GCGGCGCGCT
501 GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGCA
551 AGCCTGTTGA CCGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTTTG GGCAACATCC TTGTCGCCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
801 CGTACACGGC GGCCCGTTTC CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CCGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CTTGCGCGGT TTGAAACCTG ATGCGGCTGT TGTCGTGGCG ACTGTCCGCG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAT
1051 TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTGCGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151 TGTCGCACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACCTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCAAAAAAGT GTACGGCGCG GAAGATGTTG ATTTTCAGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCCT
1551 TTCCGAGGCG GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGACGCA
1651 GAAGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```

a141.pep
1 MSFKTDAEIA QSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
51 PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFGVKGGAA GGGYAQVLP EDINLHFTGD FHAIGAANL LAAMLNDHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPVGDVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTFAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 EGVIHGLF*

```

m141/a141 99.5% identity in 558 aa overlap

```

          10      20      30      40      50      60
m141.pep MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKLPQKQGRLLIV
          |||
a141      MSFKTDAEIAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKLPQKQGRLLIV
          10      20      30      40      50      60

          70      80      90      100     110     120
m141.pep TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGGAAAGGGYAQVLP
          |||
a141      TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGGAAAGGGYAQVLP
          70      80      90      100     110     120

          130     140     150     160     170     180
m141.pep EDINLHFTGDFHAIGAANNLLAAMLNDHIYQGNELNIDPKRVLWRRVVDNMNDRQLRNIID
          |||
a141      EDINLHFTGDFHAIGAANNLLAAMLNDHIYQGNELNIDPKRVLWRRVVDNMNDRQLRNIID
          130     140     150     160     170     180

```


m141.pep	190	200	210	220	230	240
	GMGKPVVDGVMRPDGFDTVASEVMAVFCLAKDISDLKERLGNIILVAYAKDGSPVYAKDLK					
a141	GMGKPVVDGVMRPDGFDTVASEVMAVFCLAKDISDLKERLGNIILVAYAKDGSPVYAKDLK					
	190	200	210	220	230	240
m141.pep	250	260	270	280	290	300
	ANGAMAALLKDAIKPNLVQTIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
a141	ANGAMAALLKDAIKPNLVQTIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
	250	260	270	280	290	300
m141.pep	310	320	330	340	350	360
	GFGADLGA EKFC DIKCR LAGLKPDA AVV VATVRALKYNGGVERANLGEENLDALEKGLPN					
a141	GFGADLGA EKFC DIKCR LAGLKPDA AVV VATVRALKYNGGVERANLGEENLDALEKGLPN					
	310	320	330	340	350	360
m141.pep	370	380	390	400	410	420
	LLKHISNLKNVFGLPVVVALNRFVSDADAELAMIEKACAEHGVEVSLTEVWGKGGAGGAD					
a141	LLKHISNLKNVFGLPVVVALNRFVSDADAELAMIEKACAEHGVEVSLTEVWGKGGAGGAD					
	370	380	390	400	410	420
m141.pep	430	440	450	460	470	480
	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
a141	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
	430	440	450	460	470	480
m141.pep	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
a141	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
	490	500	510	520	530	540
m141.pep	550	559				
	PAAEKIDVDAEGVIHGLFX					
a141	PAAEKIDVDAEGVIHGLFX					
	550					

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 559>:

```

g142.seq
1  ATGCGTGCCG ATTTCATGTT TGCCGACAAT ATGCCCCGTC AGGTGCGCCA
51  ACGCGCCTTC TATTTCAAGT TGTCCTCGTT TGCCGCGATG CCAAATATGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCCG CAAATGTTC
151 GGCAACATCC TGATGTTCTG CCGCCAGCAT ATTGATGCAG AGGCTGCCGT
201 TTTCGACAG GATcggaATG AttcgCGCAC TCCGGTTTAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCGGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCAC
351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC CGTCCCCTAT ATAAGAATGC TGCACACAAG
451 GCATCCCCC ATGTGCAGCA GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 560; ORF 142.ng>:

```

g142.pep
1  MRADFMFADN MPVQVRQRAF YFKLSRFAAM PNMVGKPLFG RQAGQPGKMF
51  GNILMFVRQH IDAEAAVFRQ DRNDSRTPVY AQHHGRRLVG NRRNRRHCNA
101 VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN RPLYKNAAHK
151 ASPHVQQF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 561>:

```
m142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
151 GGCAACATCC TGATGTTTCG CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCCGT AACCGGCGCG ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AAGATGCCAT CGCATCACGG AACGAAGTTT GAAAATTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
451 GCATCCCCcC ATGTGCAGCA GTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>:

```
m142.pep
1  MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
51  GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLVG NRRDRRHCA
101 VTPCRTVCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK
151 ASPHVQQF*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from *N. gonorrhoeae*:

```
m142/g142

      10      20      30      40      50      60
m142.pep MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFNGNILMFVRQR
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g142      MRADFMFADNMPVQVRQAFYFKLSRFAAMPNMVGKPLFGRQAGQPGKMFNGNILMFVRQH
          10      20      30      40      50      60

      70      80      90      100     110     120
m142.pep IDAEAAVFRQDRNDSRTPVDAQHHGRRLVGNRRDRRHCAVTPCRTVCRDDMNACRARCH
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g142      IDAEAAVFRQDRNDSRTPVYQHHGRRLVGNRRNRHCAVTPCRTVCRDDMNACRTGCH
          70      80      90      100     110     120

      130     140     150     159
m142.pep RITERSLKIFLQIRHFSPLNCPLYKNAAHKASPHVQQFX
          |||||:|||||:|||||:|||||:|||||:|||||:
g142      RITERSLKSFLQIRHFSPLNRPLYKNAAHKASPHVQQFX
          130     140     150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 563>:

```
a142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
151 GGCAACATCC TGATGTTTCG CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCCGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
451 GCACCCCCCA TGTGCAGCAG TTCTGATTCA AAAAGCCGTC GGTCCGACAT
501 TTCCGCGCGT TACGGCGTAT TACGAGTTCA ACGCATCCTC GATTTTGGCA
551 AGTTCTGCCA ACAGGTCTTT AAGCAGCAGC ATTTTCTCGC GGCCAGCAC
601 TTCTCGATA GCGTCGTAAC GCTCGTCCAC TTCTTCGCCG ATTTCCTCAT
651 ACAGCTTCTC GCCCTCGGCA GTCAGTTTCA GAAAAACACG TCGTTGGTCG
701 TTGGAAGGTT TCAGGCGGAC AACCAGCCCG GCTTTTTCAG GGCGGGTCAG
751 GATACCGGTC AGGCTGGGCG GCAAAATGCA CGCCTGATTC GCCAAATCTT
```

```

801 GAAAGTCCAG CGTGCCGTTT TCCGCCAAAA GACGGATAAT CCGCCATTGC
851 TGATCGGTAA TATTCGCCTG ATTCAGAATA GGCCTGAATT GGGTCATCAG
901 GGCTTCCCTT GCCTGTATCA GACCGATATT GATAGACGCA TGTTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 564; ORF 142.a>:

```

a142.pep
1 MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
51 GNILMFVRQR IDAAAVFRQ DRNDSRTPVD AQHHGRRLVR NRRNRHRCNA
101 VTPCRTVCRD DMNACRTGCH RITERSLSKF LQIRHFSPLN CPLYKNAAHK
151 APPMCSSTDS KSRRSDISAR YGVLRVQRIL DFGKFCQQVF KQQHFLAAQH
201 FLDSVVTLVH FFADFLIQLL ALGSQLQKNT SLVVGRFQAD NQTRFFKAGQ
251 DTGQAGAQNA RLIRQILKVQ RAVFRQKTDN PPLLIGNIRL IQNRPELGHQ
301 GFPCLYQTDI DRRMF*

```

m142/a142 96.1% identity in 153 aa overlap

```

              10      20      30      40      50      60
m142.pep    MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFNGNILMFVRQR
a142         |||||
              10      20      30      40      50      60
m142.pep    MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFNGNILMFVRQR
a142         |||||

              70      80      90      100     110     120
m142.pep    IDAAAVFRQDRNDSRTPVDAQHHGRRLVGNRRDRRHCAVTPCRTVCRDDMNACRARCH
a142         |||||
              70      80      90      100     110     120
m142.pep    IDAAAVFRQDRNDSRTPVDAQHHGRRLVNRNRNRHCAVTPCRTVCRDDMNACRTGCH
a142         |||||

              130     140     150     159
m142.pep    RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX
a142         |||||
              130     140     150     160     170     180
m142.pep    RITERSLSKIFLQIRHFSPLNCPPLYKNAAHKAPPMCSSTDSKSRRSDISARYGVLRVQRIL
a142         |||||
              130     140     150     160     170     180
m142.pep    DFGKFCQQVFKQQHFLAAQHFLDSVVTLVHFFADFLIQLLALGSQLQKNTSLVVGRFQAD
a142         |||||
              190     200     210     220     230     240

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 565>:

```

g143.seq
1 ATGTTGAGCT TCGGCTATCT CGGCGTTCAG ACGGCTTTA CCCTGCAAAG
51 CTCGCAGATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAAAT
101 TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CGGGGATGCT GGTTAGCCG
151 ATAGTGgGCT ACTACTCAGA CCGCACTTGG AAGCCGCGCT TGGGCGGCCG
201 CCGCCTGCCG TATCTGCTTT ACGGCACGCT GATTGCGGTC ATCGTGATGA
251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
301 GCCTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTGGACG TGTCGTCGAA
351 TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGATATG GTCAACGAGG
401 AGCAGAAAAG CTACGCCTAC GGGATTCAA GTTTCTTAGC GAATACGGAC
451 GCGGTTGTGG CAGCGATTCT GCCGTTTGTG TTcgcgata TCGGTTTGGC
501 GAACACTGCC GAGAAAGGCG TTGTGCCACA AACCGTGGTC GTAGATTCT
551 ATGTGGGTGC GCGGTTACTG ATTATTACCA GTGCGTTCAC AATCTCCAAA
601 GTCAAAGAAT ACGACCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
651 CGCCGCGAAT CAGGAAAAAG CCAACTGGTT CGAACTCTTA AAAACCGCGC
701 CTAAAGTGTT TTGGACGGTT ACTCCGGTAC AGTTTTTCTG CTGGTTCGCC
751 TTCCGGTATA TGTGGACTTA CTCGGCAGGC GCGATTGCAG AAAACGTCTG
801 GCACACTACC GATGCGTCTT CCGTAGGCCA TCAGGAGGCG GGCAACCGGT
851 ACGGCGTTTT GGCGGCGGTG TAGTCGGTTG CGGCGGTGAT TTGTTGGTTT
901 ATTCTGGCAA AAGTACCGAA TAAATACCAT AAGGCGGGTT ATTTGCGCTG
951 TTTGGCTTTG GGCGGCTCG GTTCTTCTC TATCTTCTC ATCTACAATC
1001 AATACGCACT CATCCTGTCT TATATCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGGCC AACGCTTGT CGGGCAAACA
1101 CATGGATACT TATTTGGGCC TGttaaagg ctctgtCTGT ATGcgcgaaa
1151 tcgTcgctTC GctgttgAGT TTCGTGCTTT TCCCGATGCT GGGCGGCCAT

```

1201 CAGGCAACCA TGTTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT
1251 CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

g143.pep
1 MLSFGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
51 IVGYSDRTW KPRLGRRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSYAY GIQSFLANTD
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL IITSFTISK
201 VKEYDPETYA RYHGIDVAAN QEKANWFELL KTAPKVFETV TPVQFFCWFA
251 FRYMWTSAG AIAENVWHTT DASSVGHQEA GNRYGVLAHV *SVAAVICSE
301 ILAKVPNKYH KAGYFGCLAL GALGFFSIFV IYNQYALILS YILIGIAWAG
351 IITYPLTIVA NALSGKHMDT YLGLFNGSVC MPQIVASLLS FVLFPMLGGH
401 QATMFLVAGA VLLLGAFSVC LIKEIHGGV*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 567>:

m143.seq
1 ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAG
51 CTCGCAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAAAT
101 TGGGCTGGTT TTTATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG
151 ATTGTCGGCC ATTATCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG
201 CCGTCTGCCG TATCTGCTTT ATGGCACGCT GATTGCGGTT ATTGTGATGA
251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
301 GCTTTGTCTG TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTCGTCAAA
351 TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG
401 AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTTCTTAGC AAATACGGGC
451 GCGGTCGTGG CGGCGATTCT GCCGTTTGTG TTTGCGTATA TCGGTTTGGC
501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTTT
551 ATGTGGGTGC GCGGTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA
601 GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
651 CGCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC
701 CTAAGGCGTT TTGGACGGTT ACTTTGGTGC AATTCTTCTG CTGGTTCGCC
751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGCTCG
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
851 ACGGCGTTTT GCGGCGGGTG CAGTCGGTTG CGGCGGTGAT TTGTTCTGTT
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTGCGCTG
951 TTTGGCTTTG GCGGCGCTCG GCTTTTTCTC CGTTTTCTTC ATCGGCAACC
1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101 TATGGGCACT TACTTGGGCT TGTTTAACGG CTCTATCTGT ATGCCTCAAA
1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201 CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251 TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

m143.pep
1 MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
51 IVGHYSRTW KPRLGRRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIQSFLANTG
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSFTIFK
201 VKEYDPETYA RYHGIDVAAN QEKANWIELL KTAPKAFETV TLVQFFCWFA
251 FQYMWTSAG AIAENVWHTT DASSVGYQEA GNWYGVLAHV QSVAAVICSE
301 VLAQVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351 IITYPLTIVT NALSGKHMTG YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
401 QATMFLVGGV VLLLGAFSVF LIKEITHGGV*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m143 / g143 93.9% identity in 429 aa overlap

10 20 30 40 50 60

```

m143.pep  MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW
g143      MLSFGYLGVTQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGYSDRTW
           10      20      30      40      50      60

           70      80      90      100     110     120
m143.pep  KPRLGGRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM
g143      KPRLGGRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM
           70      80      90      100     110     120

           130     140     150     160     170     180
m143.pep  QPFKMMVGDVNEEQKGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTAEGKGVVPQTVV
g143      QPFKMMVGDVNEEQKSYAYGIQSFLANTDAVVAAILPFVFAYIGLANTAEGKGVVPQTVV
           130     140     150     160     170     180

           190     200     210     220     230     240
m143.pep  VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTA PKAFWTV
g143      VAFYVGAALLIITSAFTISKVKEYDPETYARYHGIDVAANQEKANWFELLKTA PKVFWTV
           190     200     210     220     230     240

           250     260     270     280     290     300
m143.pep  TLVQFFCWFAFYQYMWYTSAGAI AENVWHTTDASSVG YQEAGN WYGVLA AVQSVA AVICSF
g143      TPVQFFCWFAFRYMWYTSAGAI AENVWHTTDASSVGH QEAGN RYGVLA AVXSVA AVICSF
           250     260     270     280     290     300

           310     320     330     340     350     360
m143.pep  VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT
g143      ILAKVPNKYHKAGYFGCLALGALGFFSIFFIYNQYALILSYILIGIAWAGIITYPLTIVA
           310     320     330     340     350     360

           370     380     390     400     410     420
m143.pep  NALSGKHMGTYLGLFN GSICMPQIVASLLSFVLPMLGGLQATMFLVGGV VLLGAFSVF
g143      NALSGKHMDTYLGLFN GSVCMPQIVASLLSFVLPMLGGHQATMFLVAGAVLLGAFSVC
           370     380     390     400     410     420

           430
m143.pep  LIKETHGGVX
g143      LIKEIHGGVX
           430

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 569>:

```

a143.seq
1  ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAGG
51  CTCGCAGATG AGCCGCATCT TCCAGACGCT CGGTGCCGAT CCGCACAGCC
101 TCGGCTGGTT CTTATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG
151 ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG
201 CCGTCTGCCG TATCTGCTTT ATGGCAGCCT GATTGCGGTT ATTGTGATGA
251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
301 GCTTTGTCGT TCGGCGCGCT GATGATTGCG CTGTAGACG TGTCTCAAA
351 TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG
401 AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTTCCTAGC GAATACGGGC
451 GCGGTCGTGG CGGCGATTCT GCCGTTGTG TTTGCGTATA TCGGTTTGGC
501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTT
551 ATGTGGGTGC GCGTGTGCTG GTGATTACCA GCGCGTTCAC GATTTCAAA
601 GTGAAGGAAT ACAATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
651 CGCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC
701 CTAAGGCGTT TTGGACGGTT ACTTTGGTGC AATTCTTCTG CTGGTTCGCC

```

```

751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACGGT
851 ACGGCGTTTT GCGGCGGTG CAGTCGGTTG CGGCGGTGAT TTGTTTCGTT
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTTCGGCTG
951 TTTGGCTTTG GCGGCGCTCG GCTTTTCTC CGTTTTCTC ATCGGCAACC
1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101 TATGGGCACT TACTTGGGCC TGTTTAACGG CTCTATCTGT ATGCCGCAA
1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201 CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251 TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 570; ORF 143.a>:

```

a143.pep
1  MLSEFGLGVQ TAFTLQSSQM SRIFQTLGAD PHSLGWFFIL PPLAGMLVQP
51  IVGHYSRDTW KPRLGGRRLP YLLYGTIAIV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGD VNEEQKGYAY GIQSFLANTG
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSAFTIFK
201 VKEYNPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
251 FQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAVICSF
301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351 IITYPLTI IVTNALSGKHMGT YLGFLNGSIC MPQIVASLLS FVLFPMGLGL
401 QATMFLVGGV VLLLGAFSVF LIKETHGGV*

```

m143/a143 99.5% identity in 429 aa overlap

m143.pep	10	20	30	40	50	60
a143	10	20	30	40	50	60
m143.pep	70	80	90	100	110	120
a143	70	80	90	100	110	120
m143.pep	130	140	150	160	170	180
a143	130	140	150	160	170	180
m143.pep	190	200	210	220	230	240
a143	190	200	210	220	230	240
m143.pep	250	260	270	280	290	300
a143	250	260	270	280	290	300
m143.pep	310	320	330	340	350	360
a143	310	320	330	340	350	360
m143.pep	370	380	390	400	410	420
a143	370	380	390	400	410	420

```

a143      |||||
          NALSGKHMGTYLGLFNLSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF
          370      380      390      400      410      420

          430
m143.pep  LIKETHGGVX
          |||||
a143      LIKETHGGVX
          430

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 571>:

```

g144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGGGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGTGC GTCTTCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC CCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGTAGGACTT ACCGCGTGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CCGTTCGCAC GGGCTGGCCG TTACCcgtTT
351 CAACGCGGTG GCGGCAGACG GccgaaggTt atCCCAACGA TTTGGatatt
401 TCctaccgCT TGGACGAGGA CGGCCGGCTT ACCGTtaccT ATCGCGCCAC
451 CGCgctCGGC GACACGGTGT TCGACCCGAC GCTGCACATT TACTGGCGGC
501 TGGACGCGGG CCTGCACGAT GCGGTTCTGC ATATCCGCA GGGCGGACAT
551 ATTCGGGCGG ATGCCGAAAA ACTGCCCGTC TTAACGGTTT CAGACGGCCT
601 CGAAGTATTT GA

```

This corresponds to the amino acid sequence <SEQ ID 572; ORF 144.ng>:

```

g144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTC VFVLDLGGIV QEFSVLADGV
51  RENPVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRRLSQR FGyFLPLGRG RPAYRYLSRH
151 RARRHGVRPD AAHLLAAGRG PARCGSAYSA GRYSGRCRK TARLNGFRRP
201 RSI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 573>:

```

m144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGTCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGATGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAAC AGATAGGGCG CGTGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CCGTTCGCAC GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTG GCGGCAGACG GCCGTTCGGT GGTGCTGCGC AGCCGCCTGg
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTTGG ATATTTCCTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TAaCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATGCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTCAGAC GACCTCGAAG
651 TATTGTA

```

This corresponds to the amino acid sequence <SEQ ID 574; ORF 144>:

```

m144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLATVGRRL SQRFGFYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAHLL AAGRGPARGC SAYSAGRTYA
201 GRCRK TARLN GFRRPRSI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m144 / g144 91.3% identity in 218 aa overlap

	10	20	30	40	50	60
m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD					
g144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFVLDLGGIVQEFVSLADGVRENVPVVSFDD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m144.pep	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					
g144	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m144.pep	AADGRSVVLRSLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL					
g144	AAD-----GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL					
		130	140	150	160	
	190	200	210	219		
m144.pep	AAGRGPARGCSAYSAGRTYAGRCRKRTARLNGFRRPRSI					
g144	AAGRGPARGCSAYSAGRTYSGRCRKRTARLNGFRRPRSI					
	170	180	190	200		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 575>:

```

a144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CCGTTCGCAC GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTG GCGGCAGACG GCCGTTCCGT GGTGCTGCGC AGCCGCCTG.
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTGG ATATTTCCTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
651 TATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:

```

a144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFVSLADGV
51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAHLL AAGRGPARGC SAYSAGRTYS
201 GRCKRTARLN GFRRPRSI*

```

m144/a144 99.1% identity in 218 aa overlap

	10	20	30	40	50	60
m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD					
a144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m144.pep	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					


```

|||||
a144      AASYADNPFOINKQIGRVAGRIIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
              70      80      90      100      110      120
              130      140      150      160      170      180
m144.pep  AADGRSVVLRSLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAHLL
a144      AADGRSVVLRSLXTVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAHLL
              130      140      150      160      170      180
              190      200      210      219
m144.pep  AAGRGPARGSAYSAGRTYAGRCRKRTARLNGFRFRPSIX
              |||||:|||||
a144      AAGRGPARGSAYSAGRTYSGRCRKRTARLNGFRFRPSIX
              190      200      210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 577>:

```

g146.seq
1  ATGAAGCAAA TCCCCCTCCG CCTTCTCCAG GTCGTCATTG ACCACGACAA
51 AGTCGAACAA TACGGACTGT TCGATTTCAT GCCTTGCCCT CGACAGCCTC
101 CTTTGGATAa ctTCCCGACT GTCCGTCCCG CGCctTTGA GGCGCGCGGC
151 AAGCACGTCG AAAGAAGGCG GCAGGATAAA GATACCGACA GCTTCCGGCA
201 GCGCGTTGCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACAG CCCTCCGTGC TTGTGCCGTA
301 ATAGTTGCCG AATACGTCTG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGATTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTATGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAA CCGATTGTAT CACAACGGAC ACCCTATTTT
601 ATATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 578; ORF 146.ng>:

```

g146.pep
1  MKQIPLRLQ VVIDHDKVEQ YGLFDFMPCL RQPPLDNFPT VRPAPFEARG
51 KHVERRRQDK DTDSFRQVA NLRRALNVDF QNHVIACRRQ RIHALRACAV
101 IVAEYVCVFQ KSLLRDKRFK LFFGNKVIMY AVCFAFTRRA RMRHRGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPYF
201 IFADAHILPL LF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 579>:

```

m146.seq
1  ATGGCGCAAA TCCTCCTCCG CTCGCGCCAA GTCGTCATTG ACCACGACAA
51 AGTCAAACAA TACGGACTGC TCGATTTCAT GCCTTGCCCT CGACAGCCTC
101 CTTTGGATAA CTTCCCGACT GTCCGTCCCG CGTCCGTTGA GGCGCGCGGC
151 AAGTACGTCG AAAGAAGGCG GCAGGATAAA GATGCCGACG GCTTCCGGCA
201 GCGCGTCGCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACAC CCCTCCGCGC CTGTGCCGTA
301 ATAGTTGCCA AATACGTCTG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTGTGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAATGGAC ACCCAGTTTC
601 CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 580; ORF 146>:

```

m146.pep
1  MAQILLRSRQ VVIDHDKVKQ YGLLDFMPCL RQPPLDNFPT VRPASVEARG
51 KYVERRRQDK DADGFGQVA NLRRALNVDF QNHVIACRRQ RIHTLRACAV
101 IVAKYVGVFQ KSFLRDKRLK LFFGNKVIMY AVCFAFTRRA RVRHRGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQWTPSF
201 LFADAHILPL LF*

```

Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from *N. gonorrhoeae*

m146 / g146 90.1% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARGKYVERRRQDK					
g146	MKQIPLRLQVVVIDHDKVEQYGLDFMPCLRQPPLDNFPTVRPAPFEARGKHVERRRQDK					
	10	20	30	40	50	60
	70	80	90	100	110	120
m146.pep	DADGFGQRVANLRRALNVDFQNHVIACRRQRIHTLRACAVIVAKYVGVFQKSFLRDKRLK					
	:					
g146	DTDSFRQRVANLRRALNVDFQNHVIACRRQRIHALRACAVIVAEYVCVFQKSLLRDKRFK					
	70	80	90	100	110	120
	130	140	150	160	170	180
m146.pep	LFFGNKVIMYAVCFATRRARRVRHGNQTMVCQQPRHQRFARAGSGRNDKDVAFSIS					
g146	LFFGNKVIMYAVCFATRRARRMRHGNQTMVCQQPRHQRFARAGSGRNDKDVAFSIS					
	130	140	150	160	170	180
	190	200	210			
m146.pep	GHIFYLYIFQPIVSQWTPSFLFADAHILPLLEFX					
g146	GHIFYLYIFQPIVSQRTPYFIFADAHILPLLEFX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 581>:

```

a146.seq
1  ATGGCGCAAA TCCTCCTCCG CCCGCGCCAA GTCATCATTG ACCACGACAA
51  AATCGAACAA TACGGACTGT TCGATTTCAT GCCTTGCCTT CGACAGCCTC
101 CTTTGGATAA CTCCCCGACT GTCCGTCCCG CGTCCGTGTA GACGCGCAGC
151 AAGCACATCG AAAGACGGCG GCAGGATAAA GATGCCGACG GCTTCGGGCA
201 GCGCATCTCG AACCTGAGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAACCTG CCGCCGCCAA CGCATTCACA CCCTCCGCGC TTGTGCCGTA
301 ATAGTTGCCG AACACGTCCG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGACG CGTCGTGTGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAACGGAC ACCCGGTTTC
601 CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 582; ORF 146.a>:

```

a146.pep
1  MAQILLRPRQ VIIDHDKIEQ YGLDFMPCL RQPPLDNFPT VRPASVETR
51  KHIERRRQDK DADGFGQRIS NLSRALNVDF QNHVITCRQ RIHTLRACAV
101 IVAEHVRVFQ KSLLRDKRLK LFFGNKVIMY AVCFATRRRT RVRVRHGNQ
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPGF
201 LFADAHILPL LF*

```

m146/a146 90.6% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARGKYVERRRQDK					
a146	MAQILLRPRQVIIDHDKIEQYGLDFMPCLRQPPLDNFPTVRPASVETRSKHIERRRQDK					
	10	20	30	40	50	60
	70	80	90	100	110	120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 583>:

```

g147.seq (partial)
1 ..ATGCGACGAG AAGCCAAAAT GGCACAAATC AACTCAAAC CCATTGTTTT
51 ATCAATTCTT TTAATCAACA CACCCCTCCT CGCCCAAGCG CATGAACTG
101 AGCAATCGGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151 CGCGCGACTT CGGGGCTGCT GCACACTTCG ACCGCCTCCG ACAAATCAT
201 CTCGGCGGAT ACTTTGCGCC AAAAAGCCGT CAACTTGGGC GACGCTTTGG
251 ACGGCGTACC GGGCATCCAC GCTTCGCAAT ACGGCGGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC GGGCAGACGG ATTAAAGTAT TGAACCATCA
351 CGGCGAAACG GGCGATATGG CGGACTTTTC TCCCGATCAC GCCATTATGG
401 TAGATACCGC CTTGTGCGAA CAGGTTGAAA TCCTGCGCGG GCCGTTTACG
451 CTCTTGTTACA CTTCGGGcaa tgtggcgg GCTGTGcaat gttgccgatg
501 gAAAAAtccc ccaaaaAAtg cc..

```

This corresponds to the amino acid sequence <SEQ ID 584; ORF 147.ng>:

```

g147.pep      (partial)
  1  ..MRREAKMAQI TLKPIVLSIL LINTPLLAQA HETEQSVGLE TVSVVGKSRP
51  RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGR IKVLNHHGET GDMADFSPDH AIMVDTALSQ QVEILRGPVT
151 LLYSSGNVAG AGQCCRWNKP PKNA..

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 585>:

m147.seq	(partial)					
1	..CCGCATAAAA	CTGAGCAATC	GGTGGATTTG	GAAACGGTCA	GCGTCGTCGG	
51	CAAAAGCCGT	CCGCGCGCCA	CGTCGGGGCT	GTTGCACAT	TCGCAGCCCT	
101	CCGACAAAAT	CATCTCCGGC	GATACCTTGC	GCTCAAAAAGC	CGTCAACTTG	
151	GGCGACGGTT	TAGACGCGCT	ACCGGGCATC	CACGCTTCGC	AATACGGCGG	
201	CGGCGCGTCT	GCTCCCCTCA	TTCGCGGTCA	AACAGGCAGG	CGGATTAAAG	
251	TGTTGAACCA	TCACGGCGAA	ACAGGCGATA	TGGCGGATTT	TTCGCCCGAT	
301	CACGCCATTA	TGGTAGATAC	CGCCTTGTCG	CAACAGGTCG	AAATCCTGCG	
351	CGGGCCGGTT	ACGCTCTTGT	ACAGCTCGGG	CAATGTGGCG	GGGCTGGTCG	
401	ATGTTGCCGA	TGGCAAAATC	CCCGAAAAAA	TGCCTGAAAA	CGGCGTATCG	
451	GGCGAACTCG	GATTGCGTTT	GAGCAGCGGC	AATCTGGAAA	AACTCACGCT	
501	CGGCGGCGATC	AATATCGGTT	TGGGCAAAAA	CTTTGTATTG	CACACGGAAG	
551	GGCTGTACCG	CAAACTCGGG	GATTACCGCG	TACCGCGTTA	CCGCAATCTG	
601	AAACGCCTGC	CCGACAGCCA	CGCCGATTCT	CAAACGGGCA	GCATCGGGCT	
651	GTCTTGGGTT	GGCGAAAAAG	GTTTTATCGG	CGTAGCGTAC	AGCGACCGTC	
701	GCGACCAATA	TGGTCTGCCT	GCCCACAGCC	ACGAATACGA	TGATTGCCAC	
751	GCCGACATCA	TCTGGCAAAA	GAGCTTGATT	AGCAAAACGT	ATTTACAGCT	
801	TTATCCGCAC	CTGTTGACCG	AAGAAGACAT	GATTATCGAC	AATCCGGGCT	
851	TGAGCTGCGG	CTTCCACGAC	GACGATAATG	CACACGCACA	CACCCACAGC	
901	GGCAGACCGT	GGATAGACCT	GCGCAACAAA	CGCTACGAAC	TCCGTGCCGA	
951	ATGGAAGCA	CCGTTCCTCC	GTTTGAAGC	CCTGCGCGTA	CACCTGAACC	
1001	GCAACGACTA	CGCCACGAC	GAAAAAGCAG	GCGATGTCAGT	CGAAAACCTT	
1051	TTTAACAACC	AAACGCAAAA	CGCCCGCATC	GAGTTGCGCC	ACCAACCCAT	

```

1101 AGGTCGTCTG AAAGGCAGCT GGGGCGTGCA ATATTTACAA CAAAAATCCA
1151 GTGCTTTATC TGCCATATCC GAAGCGGTTA AACAACCGAT GCTGCTTGAC
1201 AACAAAGTGC AACATTACAG CTTTTTCGGT GTAGAACAGG CAAACTGGGA
1251 CAACTTCACG CTTGAAGGAG GCGTACGCGT GGAAAAACAA AAAGCCTCCA
1301 TTCAGTACGA CAAAGCATTG ATTGATCGGG AAAACTACTA CAACCACCCC
1351 CTGCCCAGCC TCGGCGCGCA CCGCCAAACC GCCCGCTCAT TCGCACTTTC
1401 GGGCAACTGG TATTTCACGC CACAACACAA ACTCAGCCTG ACCGCCTCCC
1451 ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAACAC
1501 GTCGCCACCA ACACCTTTGA AGTCGGCAAC AAACACCTCA ACAAAGAGCG
1551 TTCCAACAAT ATCGAACTCG CGCTGGGCTA CGAAGGCGAC CGCTGGCAAT
1601 ACAATCTGGC ACTCTACCGC AACCCTTCG GTAACCTACAT TTACGCCCAA
1651 ACCTTAAACG ACGGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT
1701 GAAGCTCGTG CGCTACAACC AATCCGCGCG CGACTTCTAC GCGCGCGAAG
1751 GCGAAATCTA CTTCAAACCG ACACCGCGCT ACCGCATCGG CGTTTCCGGC
1801 GACTATGTAC GAGGCCGTCT GAAAAACCTG CCTTCCCTAC CCGGCAGAGA
1851 AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAATGCCCC
1901 CCCGTGTTCC GGCTGCGCGC CTCGGCTTCC ACCTGAAAGC CTCGCTGACC
1951 GACCGTATCG ATGCCAATTT GGACTACTAC CGCGTGTTCG CCAAAAACAA
2001 ACTCGCCCGC TACGAAACGC GCACGCCCGG ACACCATATG CTCAACCTCG
2051 GCGCAAACCTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGGTACGTC
2101 AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCACCA GCAGCTTTCT
2151 CTCTGATACG CCGCAAATGG GCCGCAGCTT TACCGGCGGC GTGAACGTGA
2201 AGTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

```

m147.pep (partial)
  1 ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL
 51 GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGDMAFSPD
101 HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS
151 GELGLRLSSG NLEKLTSGGI NIGLGKNFVL HTEGLYRKSG DYAVPRYRNL
201 KRLPDSHADS QTGSIGLSWV GEKGFIVGAY SDRRDQYGLP AHSHEYDDCH
251 ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHHTS
301 GRPWIDLRNK RYELRAEWKQ PFPGFALRV HLNRRNDYRHD EKAGDAVENF
351 FNNQTQNARI ELRHQPIGRL KGSWGVQYLQ QKSSALSAIS EAVKQPMLLD
401 NKVQHYSFFG VEQANWDNFT LEGGVRVEKQ KASIQYDKAL IDRENYNHP
451 LPDLGAHRQT ARSFALSGNW YFTPQHKLSL TASHQERLPS TQELYAHGKH
501 VATNTFEVGN KHLNKERSNN IELALGYEGD RWQYNLALYR NFRGNYIYAO
551 TLNDGRGPKS IEDDSEMMLV RYNQSGADFY GAEGEIYFKP TPRYRIGVSG
601 DYVRGRLKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAR LGFHLKASLT
651 DRIDANLDYY RVFAQNKLAR YETRTPGHM LNLGANYRRN TRYGEWNWYV
701 KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m147 / g147 92.3% identity in 142 aa overlap

```

                                10      20      30
m147.pep                      PHKTEQSVDL ETVSVVGKSR PRATSGLLHTS
                                |:||||| ||||||| |||||||
g147      MRREAKMAQITLKPIVLSILLINTPLLAQAHE TEQSVGLETVSVVGKSR PRATSGLLHTS
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m147.pep      TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
g147      TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                70      80      90      100     110     120

                                100     110     120     130     140     150
m147.pep      GDMADFSPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVADGKIPEKMPENGVS
                ||||||| ||||||| ||||||| ||||||| : |||
g147      GDMADFSPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGAGQCCRWNPKNA

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 587>:

```

a147.seq
1  ATGCGACGAG AAGCCAAAAT GGCACAAACT AACTCAAAC CCATTGTTTT
51  ATCAATTCTT TTAATCAACA CACCCCTCCT CTCCAAGCG CATGGAAC TG
101 AGCAATCAGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151 CGCGCCACTT CGGGGCTGCT GCACACTTCT ACCGCCTCCG ACAAATCAT
201 CAGCGGCGAC ACCTTGCGAC AAAAAGCCGT CAACTTGGGT GATGCTTTAG
251 ACGGCGTACC GGGCATTTCAT GCCTCGCAAT ACGGCGGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC AGGCAGACGG ATTAAGTGT TGAACCATCA
351 CGGCGAAACG GCGGACATGG CGGACTTCTC TCCAGACCAT GCAATCATGG
401 TGGACAGCGC CTTGTGCGAA CAGGTCGAAA TCCTGCGCGG TCCGGTTACG
451 CTCTTGTAAC GCTCGGGCAA TGTGGCGGGG CTGGTCGATG TTGCCGATGG
501 CAAAATCCCC GAAAAAATGC CTGAAAACGG CGTATCGGGC GAACTCGGAT
551 TCGGTTTGAG CAGCGGCAAT CTGGAAAAAC TCACGTCCGG CGGCATCAAT
601 ATCGGTTTGG GCAAAAACCT TGTATTGCAC ACGGAAGGGC TGTACCGCAA
651 ATCGGGGGAT TACGCCGTAC CGCGTTACCG CAATCTGAAA CGCCTGCCCG
701 ACAGCCACGC CGATTTCGAA ACGGGCAGCA TCGGGCTGTC TTGGGTTGGC
751 GAAAAAGGCT TTATCGGCGC AGCATAAGC GACCGTCGCG ACCAATATGG
801 TCTGCCTGCC CACAGCCACG AATACGATGA TTGCCACGCC GACATCATCT
851 GGCAAAAGAG TTTGATTAAC AAACGCTATT TGCAGCTTTA TCCGCACCTG
901 TTGACCGAAG AAGACATCGA TTACGACAAT CCGGGCTTGA GCTGCGGCTT
951 TCACGACGAC GATGATGCAC ACGCCCATGC CCACAACGGC AAACCTTGGA
1001 TAGACCTGCG CAACAAACGC TACGAACTCC GCGCCGAATG GAAGCAACCG
1051 TTCCCCGGTT TTGAAGCCCT GCGCGTACAC CTGAACCGCA ACGACTACCG
1101 CCACGACGAA AAAGCAGGCG ATGCAGTAGA AAACCTTTTT AACAAACAAA
1151 CGCAAAACGC CCGTATCGAG TTGCGCCACC AACCCATAGG CCGTCTGAAA
1201 GGCAGCTGGG GCGTGCAATA TTTGGGACAA AAATCCAGTG CTTTATCTGC
1251 CACATCCGAA GCGGTCAAAC AACCGATGCT GCTTGACAAT AAAGTGCAAC
1301 ATTACAGCTT TTTCGGTGTA GAACAGGCAA ACTGGGACAA CTTACGCTT
1351 GAAGGCGGCG TACGCGTGGA AAAACAAAAA GCCTCCATCC GCTACGACAA
1401 AGCATTGATT GATCGGGAAG ACTACTACAA CCATCCCTCG CCCGACCTCG
1451 GCGCGCACCG CCAAACCGCC CGCTCATTCT CACTTTCGGG CAACTGGTAT
1501 TTCACGCCAC AACACAAACT CAGCCTGACC GCCTCCATC AGGAACGCCT
1551 GCCGTCAACG CAAGAGCTGT ACGCACACGG CAAACACGTC GCCACCAACA
1601 CCTTTGAAGT CGGCAACAAA CACCTCAACA AAGAGCGTTC CAACAATATC
1651 GAACTCGCGC TGGGCTACGA AGGCGACCGC TGGCAATACA ATCTGGCACT
1701 CTACCGCAAC CGCTTCGGCA ACTACATTTA CGCCCAAACC TTAAACGACG
1751 GACGCGGCCC CAAATCCATC GAAGACGACA GCGAAATGAA GCTCGTGCGC
1801 TACAACCAAT CCGGTGCGGA CTTCTACGGC GCGGAAGGCG AAATCTACTT
1851 CAAACCGACA CCGCGCTACC GCATCGGCGT TTCCGGCGAC TATGTACGAG
1901 GCCGTCTGAA AAACCTGCCT TCCCTACCCG GCAGGGAAGA CGCCTACGGC
1951 AACCGCCAC TCATTGCCCA AGCCGACCAA AACGCCCTC GCGTTCGGGC
2001 TGCGCGCCTC GGCGTCCACC TGAAAGCCTC GCTGACCGAC CGCATCGATG
2051 CCAATTTGGA CTACTACCGC GTGTTCGCCC AAAACAAACT CGCCCGCTAC
2101 GAAACGCGCA CGCCCGGACA CCATATGCTC AACCTCGGCG CAAACTACCG
2151 CCGCAATACG CGCTATGGCG AGTGGAATTG GTACGTCAAA GCCGACAACC
2201 TGCTCAACCA ATCCGTTTAC GCCCACAGCA GCTTCCTCTC TGATACGCCG
2251 CAAATGGGCC GCAGCTTTAC CGGCGGCGTG AACGTGAAGT TTAA

```

This corresponds to the amino acid sequence <SEQ ID 588; ORF 147.a>:

```

a147.pep
1  MRREAKMAQT TLKPIVLSIL LINTPLLSQA HGTEQSVGLE TVSVVGKSRP
51  RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGRR IKVLNHHGET GDMADFSPDH AIMVDSALSQ QVEILRGPVT
151 LLYSSGNVAG LVDVADGKIP EKMPENGVSQ ELGLRLSSGN LEKLTSGGIN
201 IGLGKNFVLH TEGLYRKSQD YAVPRYRNLK RLPDSHADSQ TGSIGLSWVG
251 EKGFIGAAYS DRRDQYGLPA HSHEYDDCHA DIIWQKSLIN KRYLQLYPHL
301 LTEEDIDYDN PGLSCGFHDD DDAHAHAHNG KPWIDLRNKR YELRAEWKQP
351 FPGFEALRVH LNRNDYRHDE KAGDAVENFF NNQTQNARIE LRHQPIGRLE
401 GSWGVOYLQK KSSALSATSE AVKQPMMLDN KVQHYSFFGV EQANWDNFTL
451 EGGVRVEKQK ASIRYDKALI DRENYYNHPL PDLGAHRQTA RSFALSGNWY
501 FTPQHKLSTL ASHQRPLPST QELYAHGKHV ATNTFEVGNK HLNKERSNNI
551 ELALGYEGDR WQYNLALYRN RFGNYIYAQT LNDGRGPKSI EDDSEMKLVR

```

415

601 YNQGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG
651 NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY
701 ETRTPGHHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP
751 QMGRSFTGGV NVKF*

m147/a147 98.1% identity in 734 aa overlap

m147.pep				10	20	30
				PHKTEQSV	DLETVSVVGKSRPRATS	SGLLHTS
a147	MRREAKMAQTTLKPIVLSILLINTPLLSQA	HGTEQSVGLETVSVVGKSRPRATS	SGLLHTS			
	10	20	30	40	50	60
m147.pep		40	50	60	70	80
	TASDKIISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRGQTGRRIKVLNHHGET				
a147	TASDKIISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRGQTGRRIKVLNHHGET				
	70	80	90	100	110	120
m147.pep		100	110	120	130	140
	GDMADFS	PDHAIMVDTALSQQVEILRGPVTL	LYSSGNVAGLVDVADGKIPEKMPENGVS	G		
a147	GDMADFS	PDHAIMVDSALSQQVEILRGPVTL	LYSSGNVAGLVDVADGKIPEKMPENGVS	G		
	130	140	150	160	170	180
m147.pep		160	170	180	190	200
	ELGLRLSSGNLEKLTSGGINIGLGKNFV	LHTEGLYRKSGDYAVPRYRNLRPLD	SHADSQ			
a147	ELGLRLSSGNLEKLTSGGINIGLGKNFV	LHTEGLYRKSGDYAVPRYRNLRPLD	SHADSQ			
	190	200	210	220	230	240
m147.pep		220	230	240	250	260
	TGSIGLSWVGEGKFIGVAYS	DRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHL				
a147	TGSIGLSWVGEGKFIGAAYS	DRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHL				
	250	260	270	280	290	300
m147.pep		280	290	300	310	320
	LTEEDIDYDN	PGLSCGFHDDDNAHAH	THSGRPWIDLRNKRYELRAEWKQPFPGFEALRVH			
a147	LTEEDIDYDN	PGLSCGFHDDDDAHAH	ANGKFWIDLRNKRYELRAEWKQPFPGFEALRVH			
	310	320	330	340	350	360
m147.pep		340	350	360	370	380
	LNRNDYRHDEKAGDAVENFFN	QNTQNARIELRHQPIGRLKGSWGVQYLQKSSALSATSE				
a147	LNRNDYRHDEKAGDAVENFFN	QNTQNARIELRHQPIGRLKGSWGVQYLQKSSALSATSE				
	370	380	390	400	410	420
m147.pep		400	410	420	430	440
	AVKQPM	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASI	QYDKALIDRENYNHNPL			
a147	AVKQPM	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASI	RYDKALIDRENYNHNPL			
	430	440	450	460	470	480
m147.pep		460	470	480	490	500
	PDLGAHRQTARSFALSGN	WYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK				
a147	PDLGAHRQTARSFALSGN	WYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK				
	490	500	510	520	530	540
m147.pep		520	530	540	550	560
	HLNKERSN	NIELALGYEGDRWQYNLALYRN	FRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR			
a147	HLNKERSN	NIELALGYEGDRWQYNLALYRN	FRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR			

416

	550	560	570	580	590	600
m147.pep	580	590	600	610	620	630
	YNQSGADFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPPIAQDDQ					
a147	YNQSGADFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPPIAQADQ					
	610	620	630	640	650	660
m147.pep	640	650	660	670	680	690
	NAPRVPAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNT					
a147	NAPRVPAARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNT					
	670	680	690	700	710	720
m147.pep	700	710	720	730		
	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
a147	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
	730	740	750	760		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 589>:

```

g148.seq
1  ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGctgg ttcaTCCCGA
51  AgctATgagt gtccgCGCGC TTGccgAcaa AATCCGCAAA AtcgaAAact
101 gGCCGCAAAA AGgcaTCTTA TTCCACGACA TCACGCCCGT CCTGCAAAGT
151 GCGGAATACT TCCGCCTTTT GGTTCGATTG CTGGTTTACC GCTATATGGA
201 TCAGAAAATC GACATCGTTG CCGGCTTGGA CGCGCGCGGC TTCATTATCG
251 GCGCGGCACT CGCCTACCAG CTCAaCGtcg gctTCGTCCC CATCCGCAAA
301 AAAGGCAAGC TGCCTTTTGA AACCCTATCG CAAAGCTAcg cgcTCGAATA
351 CGGGGAAGCT GCGGTGGAAA TCCACACCGa tgccgTCAAA CCCGGTTCGC
401 GCGTCCTGCT GGTTCGATGAT TTGGTTGCCA CGGGCGGCAC AATGCTTGCC
451 GGGCTGGAAC TGATCCGCAA ACTCGGCGGG GAAATTGTCT AAgccgCCGC
501 CATTTTGGAA TTTACCGACC TTCAAGGCGG CAAGAATATC CGCGCAAGTG
551 GCGCGCCCTT ATTTACCCTG CTTCAAAACG AAGGCTGCAT GAAAGGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 590; ORF 148.ng>:

```

g148.pep
1  MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS
51  AEYFRLLVDL LVYRYMDQKI DIVAGLDARG FIIGAALAYQ LNVGFVPIRK
101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK PGRSVLLVDD LVATGGTMLA
151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 591>:

```

m148.seq
1  ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGCTGG TTCATCCCGA
51  AGCTATGAGT GTCGGCGCGC TTGCCGACAA AATCCGCAAA ATCGAAAAC
101 GGCCGCAAAA AGGCATCTTA TTCCACGACA TCACGCCCGT CCTCAAAGC
151 GCGGAATACT TCCGCCTTTT GGTTCGATTG TTGGTTTACC GCTATATGGA
201 TCAGAAAATC GACATCGTTG CCGGTTTGGA CGCGCGCGGC TTCATTATCG
251 GCGCGGCACT CGCCTACCAG CTCAACGTCG GTTTCGTCCC CATCCGCAAA
301 AAAGGCAAGC TGCCTTTTGA AACCCTATCG CAAAGCTACG CGCTCGAATA
351 CGGGGAAGCT GCGGTGGAAA TCCACACCGA TGCCGTCAAA CTCGGTTCGC
401 GCGTGCTGCT GGTTCGATGAT TTGATTGCCA CGGGCGGCAC GATGCTTGCC
451 GGAATGGAAC TGATCCGCAA ACTCGGCGGA GAAATTGTCT AAGCCGCCGC
501 CATTTTGGAA TTTACCGACC TTCAAGGCGG CAAGAATATC CGTGCAAGCG
551 GCGCGCCCTT ATTTACCCTG CTTCAAAACG AAGGCTGTAT GAAGGGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 592; ORF 148>:

```

m148.pep
1  MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS
51  AEYFRLLVDL LVYRYMDQKI DIVAGLDARG FIIGAALAYQ LNVGFVPIRK

```

101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA
 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m148 / g148 99.0% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
g148	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
	10	20	30	40	50	60
	70	80	90	100	110	120
m148.pep	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPVIRK	KGKLPFETVS	QSYALEYGEA
g148	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPVIRK	KGKLPFETVS	QSYALEYGEA
	70	80	90	100	110	120
	130	140	150	160	170	180
m148.pep	AVEIHTDAVK	LGSRVLLVDD	LIATGGTMLA	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI
g148	AVEIHTDAVK	PGSRVLLVDD	LVATGGTMLA	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI
	130	140	150	160	170	180
	190	200				
m148.pep	RASGAPLFTL	LQNEGCMKGX				
g148	RASGAPLFTL	LQNEGCMKGX				
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 593>:

a148.seq
 1 ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGCTGG TTCATCCCGA
 51 AGCTATGAGT GTCGGTGCGC TTGCCGACAA AATCCGCAAA ATCGAAAAC
 101 GGCCGCAAAA AGGCATCTTA TTCCACGACA TCACGCCCGT CCTGCAAAAGC
 151 GCGGAATACT TCCGACTTTT GGTGATTGTA TTGGTTTACC GCTATATGGA
 201 TCAGAAAATC GACATCGTTG CCGGTTTGA CGCGCGCGGC TTCATTATCG
 251 GCGCGGCACT CGCCTACCAG CTCAACGTCG GTTTCGTCCC CATCCGCAAA
 301 AAAGGCAAGC TGCCTTTTGA AACCGTATCG CAAAGCTACG CGCTCGAATA
 351 CGGGGAAGCT GCGGTGGAAT TCCACACCGA TGCCGTCAAA CTCGGTTTCGC
 401 GCGTGCTGCT GGTGATGAT TTGGTTGCCA CGGGCGGCAC GATGCTTGCC
 451 GGAATGGAGC TGATCCGCAA ACTCGGCGGG GAAATTGTG AAGCCGCCGC
 501 CATTTTGGAA TTTACCGACC TTCAAGGCGG CAAGAATATC CGTGCAAGCG
 551 GCGCGCCCTT ATTTACCCTG CTTCAAACG AAGGCTGTAT GAAGGGCTGA

This corresponds to the amino acid sequence <SEQ ID 594; ORF 148.a>:

a148.pep
 1 MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS
 51 AEYFRLLVDL LVYRYMDQKI DIVAGLDARG FIIGAALAYQ LNVGFPVIRK
 101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LVATGGTMLA
 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

m148/a148 99.5% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
a148	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
	10	20	30	40	50	60
	70	80	90	100	110	120

418

```

m148.pep  LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
a148      LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
              70      80      90      100     110     120

              130     140     150     160     170     180
m148.pep  AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
a148      AVEIHTDAVKLGSRVLLVDDLVATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
              130     140     150     160     170     180

              190     200
m148.pep  RASGAPLFTLLQNEGCMKGX
a148      RASGAPLFTLLQNEGCMKGX
              190     200

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 595>:

```

g149.seq
1  ATGTTGATTG ACAACAATGT CCGCCATTAC AGCTTTTTCG GTGTAGAACA
51 GGCAAATTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATCCGGTAC GACAAAGCAT TGATTGATCG AGAAACTAC
151 TACAACCAGC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 GTTCGCACTT TCGGGCAACT GGTATTTTAC GCCACACCAC AAACCTAGCC
251 TGACCGCCTC CCATCAGGAa cgCCTGCCGT CAACGCaagA actGtACgca
301 caccggcAAGC ACGtcgccac CAACACCTTT GAagtcggca acaaACACCT
351 CAACAAAGaG CgttccaacA atatcgaACT CGCGCTGGgc tAcaaaggcg
401 accGCTGGCA ATACAATCTG GCAGCCTACC GCAACCGAtT CGGCAACTAC
451 ATTTACGCCC AAACCTTAaa cgacggacgC GGCCCCAAAT CCATCgaaga
501 cgacagcgaa ATGaagCTCG TGCGCTACAA CCAATCCGGT GCCGACTTCT
551 ACGgcgcgga aggcgaaatc tACTTcaaaC CGAcACCGCG CTACCGCATC
601 GGTGTTTCCG GCGACTatgt acgaggccgT CTGAAAAACC TGCCGTCCCT
651 ACCCGGCAGG gaagatcccT AcggcAAACG TCccttcaTC GCACAAGCCG
701 ACCAAAACGC CCCCCGCATT cgggtGCGC GCCTCGGCTT CCACCTGAAA
751 ACCTCGCTAA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCCAAAC AAACCTGCCC GCTACGAAAC GCGTACGCCC GGACACCATA
851 TGCTCAACCT CGGTGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGctc aACcaatCc tTACGCCCa
951 cAGCAGCTTC CTCTCTGATA CGCCGCAAAc gGCCCGCAGC TTtgcgcGCG
1001 gcgtaAACGT GaAGTTttaA

```

This corresponds to the amino acid sequence <SEQ ID 596; ORF 149.ng>:

```

g149.pep
1  MLIDNNVRHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY
51 YNQPLPDLGA HRQTARFAL SGNWYFTPHH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YKGRWQYNL AAYRNRFGNY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVSGDYVRGR LKNLPSLPGR EDPYGKRPI AQADQNAPRI PAARLGFHLK
251 TSLTDRIDAN LDYYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FAGGVNVKF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 597>:

```

m149.seq
1  ATGCTGCTTG ACAACAAAGT GCAACATTAC AGCTTTTTCG GTGTAGAACA
51 GGCAAATGCG GACAACTTCA CGCTTGAAGG AGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATTCACTAC GACAAAGCAT TGATTGATCG GGAAACTAC
151 TACAACCACC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 ATTCGCACTT TCGGGCAACT GGTATTTTAC GCCACAACAC AAACCTAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA
301 CACGCAAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGGCG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGTAACACT

```

```

451 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TGCGCTACAA CCAATCCGGC GCCGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GGCCTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
651 ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
701 ACCAAAATGC CCCCCGTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
751 GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCAAAAC AAACCTGCCC GCTACGAAAC GCGCACGCCG GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
1001 GCGTGAACGT GAAGTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:

```

m149.pep
1  MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIYQ DKALIDRENY
51  YNHPLPDLGA HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFGNY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQNAPRV PAARLGPHLK
251 ASLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng) from *N. gonorrhoeae*:

m149/g149

m149.pep	10	20	30	40	50	60
	MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIYQ DKALIDRENY YNHPLPDLGA					
g149	MLIDNNVRHYSFFGVEQANW DNFTLEGGVR VEKQKASIRYDKALIDRENY NQPLPDLGA					
	10	20	30	40	50	60
m149.pep	70	80	90	100	110	120
	HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
g149	HRQTARFAL SGNWYFTP HHKLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
	70	80	90	100	110	120
m149.pep	130	140	150	160	170	180
	RSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
g149	RSNNIELALGYKGRWQYNLAAYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
	130	140	150	160	170	180
m149.pep	190	200	210	220	230	240
	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFI AQDDQNAPRV					
g149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDPYGRPFIAQADQNAPRI					
	190	200	210	220	230	240
m149.pep	250	260	270	280	290	300
	PAARLGPHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGAN YRRNTRYGEW					
g149	PAARLGPHLKTSLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGAN YRRNTRYGEW					
	250	260	270	280	290	300
m149.pep	310	320	330	340		
	NWYVKADNLLNQSVYAHSSF LSDTPQMGRSFTGGVNVKFX					

g149 NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFAGGVNVKFX
310 320 330 340

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 599>:

a149.seq
1 ATGCTGCTTG ACAATAAAGT GCAACATTAC AGCTTTTTCG GTGTAGAACA
51 GGCAAAGCTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
101 AAAAAAGCCTC CATCCGCTAC GACAAAGCAT TGATTGATCG GGAAACTAC
151 TACAACCATC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 ATTGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACCTAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA
301 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGGCG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGCAACTAC
451 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TGCCTACAA CCAATCCGGT GCGGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GCGGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
651 ACCCGGCAGG GAAGACGCCT ACGGCAACCG CCCACTCATT GCCCAAGCCG
701 ACCAAAACGC CCCTCGCGTT CCGGCTGCGC GCCTCGGCGT CCACCTGAAA
751 GCCTCGCTGA CCGACCGCAT CGATGCCAAT TTGGACTACT ACCGCGTGT
801 CGCCAAAAC AAACCTGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCGCA ATACGCGTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTC CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
1001 GCGTGAACGT GAAGTTTAA

This corresponds to the amino acid sequence <SEQ ID 600; ORF 149.a>:

a149.pep
1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY
51 YNHPLPDIGA HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVS GDYVRGR LKNLPSLPGR EDAYGNRPLI AQADQNAPRV PAARLGVLHK
251 ASLTDRIDAN LDYYRVFAQN KLARYETRTF GHMMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

m149/a149 98.8% identity in 339 aa overlap

	10	20	30	40	50	60
m149.pep	MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIYQDKALIDRENYNHPLPDIGA					
a149	MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPLPDIGA					
	10	20	30	40	50	60
m149.pep	HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
a149	HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	70	80	90	100	110	120
m149.pep	HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
a149	HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	70	80	90	100	110	120
m149.pep	RSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
a149	RSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
	130	140	150	160	170	180
m149.pep	RSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
a149	RSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
	130	140	150	160	170	180
m149.pep	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDAYGNRPFIAQDDQNAPRV					
a149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDAYGNRPLIAQADQNAPRV					
	190	200	210	220	230	240
m149.pep	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDAYGNRPFIAQDDQNAPRV					
a149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDAYGNRPLIAQADQNAPRV					
	190	200	210	220	230	240
m149.pep	PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEW					
a149	PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEW					
	250	260	270	280	290	300
m149.pep	PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEW					
a149	PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEW					

g149-1.seq

1	ATGGCGACAAA	TCACACTCAA	ACCCATTGTT	TTATCAATTC	TTTTAATCAA
51	CACACCCCTC	CTCGCCCAAG	CGCATGAAAC	TGAGCAATCG	GTGGGGCTGG
101	AAACCGGCTAG	CGTCGTCCGG	AAAGCGGATC	CGCGCGCGAC	TCGGGGGTCT
151	CTGCACACTT	CAGCGCGCTC	CGACAAAATC	CATCTCGGCG	ATACTTTGGC
201	CCAAAAAGCC	GTCAACTTGG	GCGACGCTTT	GGACGGCGTA	CCGGGCATCC
251	ACGCTTCCGA	ATACGGCGGC	GGCGCATCCG	CTCCCGTTAT	TCGCGGTCAA
301	ACGGGACGAG	GATTAAGT	ATTGAACCAT	CACGGCGAAA	CGGGCGATAT
351	GGCGGACTTT	TCTCCCGATC	ACGCCATTAT	GGTAGATACC	CGCTTGTCTG
401	AACAGGTTGA	AATCCTCGCG	GGCGCGGTTA	CGCTCTTGTA	CAGCTAGGAT
451	AATGTGGCGG	GCGTGGTGTA	TGGTGGCGAT	GGAAAAATCC	CCGAAAAAAT
501	CCCTGAAAAA	GGCGTATCGG	GCGAagccgG	ATTGCGTTTG	AGCAGCGGCA
551	ATTTAGAAAA	ACTGACATCC	GCAGGCATCA	ATATCGCAAT	GGGCAAAAAC
601	TTCTGTCTGC	ATACCCGAAG	CTTGTAACCG	AAATCGGGCG	ATTACGCGCT
651	ACCGCGTTAC	CGCAATCTGA	ACAGCCTGCC	CGACAGCCAT	GCCGATTCTG
701	AAACGGGACG	CATCGGGCTG	TCTTGGGTGG	GCGAAAAAGG	CTTTATCTGG
751	GCAGCATACA	GCGACCTGCG	CGACCGCTAC	GGCCTGCCTG	CCCCAGCCCA
801	CGAATACGAT	TGTTGCCAC	CGCATCATCT	CTGGCAAAA	AGGTTTGATCA
851	ACAAACGCTA	TTTGACGTT	TATCTGCAT	TGTTGACCGA	AGAAGACATC
901	GATTACGACA	ATCCGGGCTT	GAGCTGCGGC	TTCACACGAC	GCGACGGTGC
951	ACACGACAC	ACCCACAACG	GCAAAACCTG	GATAGACCTC	CGCAACAAAC
1001	GCTACGAACT	CCGCGCCGAA	TGGAAGCAGC	CATTCCCCGG	TTTTGAAGCG
1051	CTGCGCGTAC	ATCTGAACCG	CAATGACTAC	CACCACGACG	AAAAAGCAGG
1101	CGATCGACGA	GAAAACTTCT	TCAACAACAA	AACACACAAC	GCCGCTATCG
1151	AGTTGCGCCA	CCAACCCATA	GGCCGTCTGA	AAGGCAGCTG	GGGCGTGCAA
1201	TATTTGGGAC	AAAAATCCAG	CGCGCTTCCG	GCCATTCCCG	AAACCGTCCA
1251	ACAACCGATG	TTGATTGACA	ACAATTCCCG	CCATTACAGC	TTTCTCGGTG
1301	TAGAACAGGC	AATTTGGGAC	AAACTTCACG	TTGAAGGCGG	CGTACGCGTG
1351	GAAAAACAAA	AAGCTCTCAT	CCGGTACGAC	AAAGCATTTA	TTGATCCGAA
1401	AAACTACTAC	AACCCAGCCC	TGCCCGACCT	CGGC CGCGAC	CGCCAAACCG
1451	CCCGCTCGTT	CGCATTTTCG	GGCAACTGGT	ATTTACAGCC	ACACCAACAA
1501	CTCAGCCTGA	CCGCTCCCA	TCAGGAACGC	CTGCGGTCAA	CGCAAGAACT
1551	GTACGCACAC	GGCAAGCAGC	TCGCCACCAA	CACCTTTGAA	GTCGGCAACA
1601	AACACCTCAA	CAAGAGCGGT	TCCAACAATA	TCGAATCCGC	GCTGGGCTAC
1651	GAAGGCGACC	GCTGGCAATA	CAATCTGGCA	GCCTACCGCA	ACCGATTCTG
1701	CAACTACATT	TACGCCCAA	CCTTAAACGA	CGGACGCGGC	CCCAAAATCCA
1751	TCGAAGACGA	CAGCGAAATG	AGBCTCGTGC	GCTACAACCA	ATCCGGTGCC
1801	GACTTCTACG	GCGCGGAAGG	CGAAATCTAC	TTCAAACCGA	CACCGCGCTA
1851	CCGATCTCGT	GTTTCCGGCG	ACTATGTACG	AGGCCGTCTG	AAAAACCTCG
1901	CGTCCCTACC	CGGCAGGAA	GATCCCTACG	GCAAACTGCT	CTTCATCGCA
1951	CAAGCCGACC	AAAACGCCCC	CCGCATTCCG	GCTGCGCGCC	TCGGCTTCCA
2001	CCTGAAAACC	TCGCTAACCG	ACCGTATCGA	TGCCAATTGG	GACTACTACT
2051	GCGTGTTCGC	CCAAAAACAA	CTCGCCCGCT	ACGAAACCTG	TAGCCCGGGA
2101	ACCATATATG	TCAACCTCGG	TGCAAACTAC	CGCGGCAATA	CGCGCTATGG
2151	CGAGTGGAA	TGGTACGTGA	AAGCCGACAA	CCTGCTCAAC	CAATCCGTTT
2201	ACGCCACACG	CAGCTTCTCT	TCTGATACCG	CGCAAAATGG	CCGCAGCTTT
2251	ACCGCGGGCG	TAAACGTGAA	GTTTTAA		

g149-1.pap

1	MAQITLTKPIV	LSILLINTPL	LAQAHETEQS	VGLETVSVVG	KSRPRATSGL
51	LHTSTASDKI	ISGDTLRQKA	VNLGDALDGV	PGIHASQYGG	GASAPLYRGG
101	TGRRRIKVLNH	HGETGDMADF	SPDHAIMDVT	ALSQSQVEILR	GPVTLTYRSG
151	NVAGLVVDAD	GKIKEKMPEN	GVSGEAGLRL	SSGNLEKITS	AGTNIGLGKN
201	FVLHTEGLYR	KSGDYAVPRY	RNLKRLPDSH	ADSQTSRISL	SWVGKGFGIF
251	AAYSDDRDRY	GLPAHSHEYD	DCHADIWQK	SILNKRYLQL	YPHLLTEEDI
301	YDYNPGLSCG	FHDGGAHAH	THNGKPWIDL	RNKRYELRAE	WKQPPFGFEA

```

351 LRVHLNRNDY HHDEKAGDAV ENFFNNKTHN ARIELRHQPI GRLKGSWGVQ
401 YLGQKSSALS AIPETVQQPM LIDNNVRHYS FFGVEQANWD NFTLEGGVRV
451 EKQKASIRYD KALIDRENY NQPLPDLGAH RQTARSFALS GNWYFTPHHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551 EGDWRQYNLA AYRNRFNGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601 DFYGAEGEII FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DPYGRPFIA
651 QADQNAPRIP AARLGFLHKT SLTDRIDANL DYYRVFAQNK LARYETRTPG
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751 TGGVNVKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 603>:

m149-1.seq

```

1 ATGGCACAAA CTACACTCAA ACCCATGTGT TTATCAATTC TTTTAATCAA
51 CACACCCCTC CTCGCCCAAG CGCATGAAAC TGAGCAATCG GTGGATTGG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCCAC GTCGGGGCTG
151 TTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACCTTGCG
201 CCAAAAAGCC GTCAACTTGG GCGACGCTTT AGACGGCGTA CCGGGCATCC
251 ACGCTTCGCA ATACGGCGGC GCGCGCTCTG CTCCCGTCAT TCGCGGTCAA
301 ACAGGCAGGC GGATTAAAGT GTTGAACCAT CACGGCGAAA CAGGCGATAT
351 GGCGGATTTT TCGCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTCGC
401 AACAGGTCGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
451 AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGCAAAATCC CCGAAAAAAT
501 GCCTGAAAAC GCGGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
551 ATCTGAAAAA ACTCACGTCC GCGGCGCATCA ATATCGGTTT GGGCAAAAAC
601 TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCCGT
651 ACCGCGTTAC CGCAATCTGA AACGCCCTGCC CGACAGCCAC GCCGATTCGC
701 AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAGG TTTTATCGGC
751 GTAGCGTACA GCGACCGTCG CGACCAATAT GGTCTGCGTG CCCACAGCCA
801 CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGCTTGATTA
851 ACAAACGCTA TTTACAGCTT TATCCGCACC TGTTGACCGA AGAAGACATC
901 GATTACGACA ATCCGGGGCTT GAGCTGCGGC TTCCACGACG ACGATAATGC
951 ACACGCACAC ACCCACAGCG GCAGACCGTG GATAGACCTG CGCAACAAAC
1001 GCTACGAACT CCGTGCCGAA TGGAAGCAAC CGTTCGCCGG TTTGAAGCC
1051 CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101 CGATGCAGTC GAAACTTTT TTAACAACCA AACGCAAAAC GCCCGCATCG
1151 AGTTGCCGCA CCAACCCATA GGTCGTCTGA AAGGCAGCTG GGGCGTGCAA
1201 TATTTACAAC AAAAATCCAG TGCTTTATCT GCCATATCCG AAGCGGTTAA
1251 ACAACCGATG CTGCTTGACA ACAAAGTGCA ACATTACAGC TTTTTCGGTG
1301 TAGAACAGGC AAACCTGGAC AACTTCACGC TTGAAGGAGG CGTACGCGTG
1351 GAAAAACAAA AAGCCTCCAT TCAGTACGAC AAAGCATTGA TTGATCGGGA
1401 AAATACTTAC AACCACCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1451 CCGGCTCATT CGCACTTTCG GGCAACTGGT ATTTCACGCC ACAACACAAA
1501 CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAGCT
1551 GTACGCACAC GGCAAAACAG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
1651 GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 TAAGTACATT TACGCCCAA CTTAAACGA CGGACGCGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGCGCC
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851 CCGCATCGGC GTTTCGGCG ACTATGTACG AGGCCGTCTG AAAAACCTGC
1901 CTTCCCTACC CGGCAGAGAA GATGCCTACG GCAACCGTCC TTTCATCGCA
1951 CAGGACGACC AAAATGCCCC CCGTGTTCCG GCTGCGCGCC TCGGCTTCCA
2001 CCTGAAAGCC TCGCTGACCG ACCGTATCGA TGCCAATTG GACTACTACC
2051 GCGTGTTTCG CCAAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2101 CACCATATGC TCAACCTCGG CGCAAACCTAC CGCCGCAATA CGCGCTATGC
2151 CGAGTGGAAAT TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2201 ACGCCACAGC CAGCTTTCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
2251 ACCGGCGGCG TGAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 604; ORF 149-1>:

m149-1.pep

```

1 MAQTLKPIV LSILLINTPL LAQAHETEQS VDLETVSVVG KSRPRATSGL
51 LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRIKVLVNH HGETGDMADF SPDHAIMVDI ALSQQVEILR GPVTLLEYSSG
151 NVAGLVVDVAD KGIPEKMPEN GVSGLGLRL SSGNLEKLT SGINIGLGKN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIF
251 VAYSDRRDQY GLPAHSHEYD DCHADIIWQK SLINKRYLQL YPHLLTEEDI
301 LDYDNPGLSCG FHDDDNAHAH THSGRPWIDL RNKRYELRAE WKQFPFGFEA
351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTON ARIELRHQPI GRLKGSWGVQ
401 YLQKSSALS AISEAVKQPM LLDNKVOHYS FFGVEQANWD NFTLEGGVRV
451 EKQKASIQYD KALIDRENY NHPLPDLGAH RQTARSFALS GNWYFTPQHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY

```

551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
 601 DFYGAEGEIIY FKPTPRYRIG VSGDYVRGRL KNLPSPGREG DAYGNRPFIA
 651 QDDQNAPRVP AARLGFHLKA SLTDRIANL DYRVFAQNK LARYETRTPG
 701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
 751 TGGVNVKF*

m149-1/g149-1 96.2% identity in 758 aa overlap

	10	20	30	40	50	60
m149-1.pep	MAQTTLKPIVLSILLINTPLLAQAHETE	QSV	VDLET	VS	VVGKSRPRAT	SGLLHTSTASDKI
g149-1	MAQITLKPIVLSILLINTPLLAQAHETE	QSV	GLET	VS	VVGKSRPRAT	SGLLHTSTASDKI
	10	20	30	40	50	60
	70	80	90	100	110	120
m149-1.pep	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
g149-1	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
	70	80	90	100	110	120
	130	140	150	160	170	180
m149-1.pep	SPDHAIMVDTALSQQVEILRGVPTLLYSSGNVAGLVDVADGKIPEKMPENGVS	GELGLRL				
g149-1	SPDHAIMVDTALSQQVEILRGVPTLLYSSGNVAGLVDVADGKIPEKMPENGVS	GEAGLRL				
	130	140	150	160	170	180
	190	200	210	220	230	240
m149-1.pep	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHAD	SQTGSIGL				
g149-1	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHAD	SQTGSIGL				
	190	200	210	220	230	240
	250	260	270	280	290	300
m149-1.pep	SWVGEKGFIVGAYSDDRDQYGLPAHSHEYDDCHADI IWQKSLINKRYLQLYPHLLTEEDI					
g149-1	SWVGEKGFIVGAYSDDRDQYGLPAHSHEYDDCHADI IWQKSLINKRYLQLYPHLLTEEDI					
	250	260	270	280	290	300
	310	320	330	340	350	360
m149-1.pep	DYDNPGLSGCFHDDDDNAHAHTHSGRPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
g149-1	DYDNPGLSGCFHDDDDNAHAHTHSGRPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
	310	320	330	340	350	360
	370	380	390	400	410	420
m149-1.pep	RHDEKAGDAVENFFNNTQNARIELRHQPIGRKGSWGVQYLQKSSALS	AISEAVKQPM				
g149-1	HHDEKAGDAVENFFNNTQNARIELRHQPIGRKGSWGVQYLQKSSALS	AIPEVQQPM				
	370	380	390	400	410	420
	430	440	450	460	470	480
m149-1.pep	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASI QYDKALIDRENYNHP	PLDLAGH				
g149-1	LIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNQ	PLDLAGH				
	430	440	450	460	470	480
	490	500	510	520	530	540
m149-1.pep	RQTARSAFALSGNWFYTPQHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLN	KER				
g149-1	RQTARSAFALSGNWFYTPHHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLN	KER				
	490	500	510	520	530	540
	550	560	570	580	590	600
m149-1.pep	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEM	KLVRYNQSGA				
g149-1	SNNIELALGYEGDRWQYNLAAYRNRFGNYIYAQTLNDGRGPKSIEDDSEM	KLVRYNQSGA				
	550	560	570	580	590	600
	610	620	630	640	650	660
m149-1.pep	DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSPGREGDAYGNRPFIAQDDQNAPRVP					
g149-1	DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSPGREDPYGKRPFIAQADQNAPRIP					
	610	620	630	640	650	660

a149-1.seq

This corresponds to the amino acid sequence <SEQ ID 606; ORF 149-1.a>:

a149-1.pgp

1	MAQTTLKP	IV	LSILLINTPL	LSQAHGTEQS	VGLETVS	VVG	KSRPRATSG	L
51	LHTSTASDKI		ISGDTLRQKA	VNLGDALDGV	PGIHASQYGG		GAPAPVIRGQ	
101	TGRRIKVNLN		HGETGDMADF	SPDHAIMDVS	ALSQSQVEILR		GSPTLLYSSG	
151	NVAGLVLDVAD		GKIPKMPEN	GVSGELGLRL	SSGNLEKITS		GGINIGLGN	
201	FVLHTEGRLY		KSGDYAVPRY	RNLKRLPDSH	ADSGTSGYL		SWGEGKFGI	
251	AAYSDDRLDQY		GLPAHSHEYD	DCHADIIWQK	SILNKRYLQL		YKPHLLTEEDI	
301	DYDNPGLSCG		FHDDDDAHAA	AHNGKPWIDL	RNKRYELRAE		WKQFPGFEA	

```

351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTQN ARIELRHQPI GRLKGSWGVQ
401 YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV
451 EKQKASIRYD KALIDRENY NHPLPDLGAH RQTARSFALS GNWYFTPOHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601 DFYGAEGEYI FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPLIA
651 QADQNAPRVP AARLGVHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751 TGGVNVKF*

```

a149-1/m149-1 98.0% identity in 758 aa overlap

	10	20	30	40	50	60
a149-1.pep	MAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI					
m149-1	MAQTTLKPIVLSILLINTPLLAQAHETEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI					
	10	20	30	40	50	60
	70	80	90	100	110	120
a149-1.pep	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
m149-1	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
	70	80	90	100	110	120
	130	140	150	160	170	180
a149-1.pep	SPDHAIMVDSALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
m149-1	SPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
	130	140	150	160	170	180
	190	200	210	220	230	240
a149-1.pep	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQTGSIGL					
m149-1	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQTGSIGL					
	190	200	210	220	230	240
	250	260	270	280	290	300
a149-1.pep	SWVGEKGFIGAAYSDDRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
m149-1	SWVGEKGFIVAYSDDRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
	250	260	270	280	290	300
	310	320	330	340	350	360
a149-1.pep	DYDNPLGLSCGFHDDDDAHAAHNGKWPWIDLRNKRYELRAEWKQFPFGFEALRVHLNRNDY					
m149-1	DYDNPLGLSCGFHDDDDAHAAHNGKWPWIDLRNKRYELRAEWKQFPFGFEALRVHLNRNDY					
	310	320	330	340	350	360
	370	380	390	400	410	420
a149-1.pep	RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSEAVKQPM					
m149-1	RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSEAVKQPM					
	370	380	390	400	410	420
	430	440	450	460	470	480
a149-1.pep	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPLPDLGAH					
m149-1	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPLPDLGAH					
	430	440	450	460	470	480
	490	500	510	520	530	540
a149-1.pep	RQTARSFALSGNWFYTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER					
m149-1	RQTARSFALSGNWFYTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER					
	490	500	510	520	530	540
	550	560	570	580	590	600
a149-1.pep	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMCLVRYNQSGA					
m149-1	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMCLVRYNQSGA					
	550	560	570	580	590	600
	610	620	630	640	650	660


```

a149-1.pep  DFYGAEGEIIYFKPTPRYRIGVSGDYVRGLKNLPSLPGREDAYGNRPILIAQADQNAPRVP
|||||
m149-1      DFYGAEGEIIYFKPTPRYRIGVSGDYVRGLKNLPSLPGREDAYGNRPFIAQDDQNAPRVP
              610      620      630      640      650      660

              670      680      690      700      710      720
a149-1.pep  AARLGVLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEWN
|||||
m149-1      AARLGVLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEWN
              670      680      690      700      710      720

              730      740      750      759
a149-1.pep  WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
|||||
m149-1      WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
              730      740      750

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 607>:

```

g150.seq (partial)
1  ..TACTGCAAGG CAGACCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAAT
51  CACCGCCCGC CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTGTA
101 GCGGTTCCGA TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT
151 GACAACGATC CGGCACTGGT CGGGGAAATC CTAGACCTGC TCGGCATCAA
201 TCCGCAACG GAAATACAGG CGGGCGGAAA AACCTGCGG GTTGCTCCG
251 CACTGTTATC CCATTTCGAA CTCACGCAA ACACCCCGC CTTTGTCAA
301 GGCTATGCCA CGTTCGCCGA TAATGACGAA CTCGACCGTA TTGCTGCCGA
351 CAACGCCGTT TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCGGTGTGC
401 TGCACCGCTT CCCGGCAAAA CTGACGGCGG AACAAATCGC CGGCCTGCTG
451 GCCTCGCTTG CGCCGCGCCT GTATTTCGATT TCCTCGTCGC AGGCGGAAGC
501 GGGGGACGAA GTGCACCTGA CCGTCGGCGC AGTGCCTTTC GAACACGAAG
551 GGCAGCGCCAG GCGGGCGGCG GCATCGGGTT TCTTTGCCGA CCGGCTGGAA
601 GAGGACGGCA CGGTGCGCGT GTTTGCGGAA CGCAACGACG GCTTCAGGCT
651 GCCCGAAGAC AGCCGCAAGC CGATTGTGAT GATCGGCTCC GGTACCGGCG
701 TCGCACCGTT CCGCGCCTTC GTCCAACAAC GTGCCGAGA AAATGCGGAA
751 GGCAGAAACT GGCTGATTTT CGGCAATCCG CATTTTGCCG CCGACTTCCT
801 CTATCAGACC GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT
851 ATGACTTCGC CTGGTCGCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC
901 AAAATCCGCG AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC
951 GCATATCTAT GTGTGCGGCG ATGCGGCAA AATGGCAAAA GAAGTGAAG
1001 CCGCCTTGCT GGATGTGATT ATCGGGGCG GGCATTCCGA CGAAGACGGC
1051 GCAGAAGGAT ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA
1101 TGTTTATTGA

```

This corresponds to the amino acid sequence <SEQ ID 608; ORF 150.ng>:

```

g150.pep (partial)
1  ..YKADPFPA LLANQKITAR QSDKDVRHIE IDLSGSDLHY LPGDALGVWF
51  DNDPALVGEI LDLLGINPAT EIQAGGKTL P VASALLSHFE LTQNTPAFVK
101 GYATFADNDE LDRIAADNAV LQGFVQSTPI AGVLHRFPAK LTAEQFAGLL
151 RPLAPRLYSI SSSQAEAGDE VHLTVGAVRF EHEGRARAGG ASGFFADRLE
201 EDGTVRVFAE RNDGFRLPED SRKPIMIGS GTGVAPFRAF VQQRAAENAE
251 GRNWLIFGNP HFAADFLYQT EWQQFAKDG F LHRDYFAWSR DQEEKIYVQD
301 KIREQAEGLW QWLQEGAHY VCGDAAKMAK EVEAALLDVI IGAGHSDG
351 AEGYLDMLRE EKRYQRDVY*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 609>:

```

m150.seq
1  ATGCAGAACA CAAATCCGCC ATTACCGCCT CTGCCGCCCG AAATCACGCA
51  GCTCCTGTCT GGGCTGGACG CGGCACAATG GGCGTGGCTG TCCGGCTACG
101 CTTGGGCAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
151 ACGGCATTGC CGGCGGCAGA ACCTTTTTC GTAACCGTCC TTTCGCCTC
201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGCG GACAGCCTGG
251 AAGCCGCCCG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
301 AAAAACATCG CCGGCGAAGC CCGCCTGCTG CTGGTTACCT CCACCCAAGG
351 CGAAGGCGAA CCGCGGAAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
401 GCAAAAAAGC CCCGAAATTG GACAACTCC AATTGCGCGT ACTGGGTTTG

```

m150.pep

1	MQNTNPPLPP	LPPEITQLLS	GLDAAQWAWL	SGYAWAKAGN	GASAGLPAQ
51	TALPAAEPPS	VTVLSASQTG	NAKSVDKAA	DSLEAAGIQV	SRAELKDYKA
101	KNIAGERLL	LVTSTQGE	PPKEAVVLHK	LLNGKKAPKL	DKLQFAVLGL
151	GDSSYPNFC	AGKDFDRRFE	ELGAKRLLER	VDAADLFTAS	ANAWTDNIAA
201	LLKEEAANKR	ATPAPOQTFP	AGLOTADPGR	YCKAADFFAA	LLANQKITAR
251	QSDKDVRIE	IDLSGSDLHY	LPGDALGVWF	DNDPALVREI	LDLLGIDPAT
301	EIQAGGKMMP	VARALSSHFE	LTQNTPAFVK	GYAAFAHYEE	LDKIIADNAV
351	LQDFVQNTPI	VDVLHRFPAS	LTAEQFIRLL	RPLAPRLYSI	SSAQAEVGDE
401	VHLTGVGVRF	EHEGRARTGG	ASGFLADRLE	EDGTVRVFE	RNDGFRFLPD
451	SRKPIVMIGS	GTGVAPFRAF	VQQRAAENAE	GKNWLIFGNP	HFARDFLPYQ
501	EWQQFKAQDG	LHRYDFAWSR	DQEEKIYVQD	KIREQAEGWL	QWLQEGAHYI
551	VCGDAAKMAK	DVEAALLDVI	IGAGHLDEEG	AEYELDMLRE	EKRYQRDVY*

Homology with a predicted ORF from *N.gonorrhoeae*

m150/q150

	210	220	230	240	250	260
m150.pep	LLKEEAAKNRATPAPQTTPAGLQTAPDGRYCKA	AAPPFAALLANQKITARQSDKDVRHIE				
g150				YCKADP	FAALLANQKITARQSDKDVRHIE	
				10	20	30
	270	280	290	300	310	320
m150.pep	IDLSGSDLHYLPGDALGVWFDNDPALVREILDLLG	IDPATEIQAGGKMMPVARALSSHFE				
g150	IDLSGSDLHYLPGDALGVWFDNDPALVGEILDLLG	INPATEIQAGGKTLPVASALLSHFE				
	40	50	60	70	80	90
	330	340	350	360	370	380
m150.pep	LTQNTPAFVKGYAAFAHYEELDKIIADNAV	LQDFVQNTPIVDVLHRFPASLTAEQFIRLL				

```

-----
a150.seq
-----

```

a150.seq					
1	ATGCAGAACA	CAAATCCGCC	ATTACCGCCT	ATGCCGCCCG	AAATCACGCA
51	GCTCCTGTCG	GGGCTGGACG	CGGCACAATG	GGCGTGGCTG	TCCGGCTACG
101	CTTGGGCAAA	AGCAGGAAAC	GGGGCATCTG	CAGGACTGCC	CGCGTTTCAG
151	ACGGCATTTG	CGACGGCAGA	ACCTTTTTCC	GTAACCGTCC	TTTCCGCCTC
201	GCAAACCGGC	AATGCGAAAT	CCGTTGCCGA	CAAAGCGGCG	GACAGCCTGG
251	AAGCCGCCGG	CATCCAAGTC	AGTCGCGCCG	AACTGAAAGA	CTATAAGGCG
301	AAAAACATCG	CCGGCGAACG	CCGCCTGCTG	CTGGTTACCT	CCACCCAAGG
351	CGAAGGCGAA	CCGCCGGAAG	AAGCCGTCGT	GCTGCACAAA	CTGCTGAACG
401	GCAAAAAAGC	CCCGAAATTG	GACAACTCC	AATTTGCCGT	ACTGGGTTTG
451	GGCGACAGCT	CCTATCCGAA	TTTCTGCCGG	GCGGGCAAAG	ATTTTCGACAA
501	ACGTTTTGAA	GAATTGGGCG	CAAAACGCCT	GCTCGAACGC	GTTGATGCGG
551	ATTTGGAATT	TGCCGCCGCC	GCAGACGGAT	GGACAGATAA	TATCGCCGCA
601	CTCTTAAAAA	AAGAAGCCCG	AAAAAACCGG	GCAACGCCCG	CGCCGCAGAC
651	AACGCCCCCC	CGCGGCCTTC	AGACGGCACC	GATATGGCAGG	TACTGCAAGG
701	CAGACCCCTT	TCCGCCCGCC	CTGCTGGCCA	ATCAGAAAAT	CACCGCCCGC
751	CAATCCGATA	AAGACGTGCG	CCACATCGAA	ATCGATTTGA	GCGGTTCTGGA
801	TTTGCCTACT	CTCCCGGGCG	ACGCGCTCGG	CGTTTGGTTT	GACAACGATC
851	CGGCACTGGT	CAGGGAATC	CTAGACCTGC	TGGCATCTGA	TCAGGCAACG
901	GAAATACAGG	CGGGCGGAAA	AACCCTGCCG	GTTGCCTCCG	CACTGTTATC
951	CCATTTTGAA	CTCACGCAAA	ACACCCCGCG	CTTTGTCAAA	GGCTATGCCC
1001	CGTTCGCCGA	TGATGACGAA	CTCGACCGTA	TTGCTGCCGA	CAACGCCGTT
1051	TTGCAAGGCT	TTGTGCAAAG	CACGCCGATT	GCCGATGTGC	TGCACCGCTT
1101	CCCGGCAAAA	CTGACAGCGG	AACAATTCCG	CGGCCTACTG	CGCCCGCTTG
1151	CGCCGCGCCT	GTATTCGATT	TCCTCGTCGC	AGGCGGAAGT	GGGGGACGAA
1201	GTGCACCTGA	CCGTCGGCGC	GGTGC GTTTC	GAACACGAAG	GGCGCGCCAG
1251	GGCGGGCGGC	GCATCGGGTT	TCCTTGCCGA	CCGGCTGGAA	GAGGACGGCA
1301	CGGTGCGCGT	GTTTGTGGAA	CGCAACGACG	GCTTCAGGCT	GCCCCAAGAC
1351	AGCCGCAAGC	CGATTGTGAT	GATCGGCTCG	GGACCCGCGC	TGCACCGGTT
1401	CCGCGCTTTC	GTCCAACAAC	GTGCCGCAGA	GTGCCGGGAA	GCGAAAAACT
1451	GGCTGTTTTT	CGGCAATCCG	CATTTTGCCC	GTGATTTTCT	CTATCAAACC
1501	GAATGGCAGC	AGTTTGCCAA	AGACGGCTTC	CTGCACAGAT	ACGATTTTCG
1551	CTGGTCGCGC	GATCAGGAAG	AAAAAATCTA	TGTGCAGGAC	AAAAATCCGCG
1601	AACAGGCGGA	AGGACTTTGG	CAATGGCTGC	AGGAAGGCGC	GCATATCTAT
1651	GTGTGCGGCG	ATGCGGCAAA	AATGGCAAAA	GACGTGGAAG	CGGCCTTGCT

1701 GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT
 1751 ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA

This corresponds to the amino acid sequence <SEQ ID 612; ORF 150.a>:

a150.pep
 1 MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
 51 TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
 101 KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGL
 151 GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAA
 201 LLKEEAAKNR ATPAPQTPP AGLQTAPDGR YCKADFPFAA LLANQKITAR
 251 QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDQAT
 301 EIQAGGKTLP VASALLSHFE LTQNTPAFVK GYAPFADDDE LDRIAADNAV
 351 LQGFVQSTPI ADVLHRFPK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE
 401 VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED
 451 SRKPIVMIGS GTGVAPFRAF VQRAAENAE GKNWLFNGP HFARDFLYQT
 501 EWQFQAKDGF LHRYPFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHY
 551 VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY*

m150/a150 94.8% identity in 599 aa overlap

	10	20	30	40	50	60
m150.pep	MQNTNPPLPP	MPPEITQLLS	GLDAAQWAWL	SGYAWAKAGN	GASAGLPALQ	TALPAAEPFS
a150	MQNTNPPLPP	MPPEITQLLS	GLDAAQWAWL	SGYAWAKAGN	GASAGLPALQ	TALPAAEPFS
	10	20	30	40	50	60
	70	80	90	100	110	120
m150.pep	VTVLSASQTG	NAKSVADKAA	DSLEAAGIQV	SRAELKDYKA	KNIAGERRLL	LVTSTQGEGE
a150	VTVLSASQTG	NAKSVADKAA	DSLEAAGIQV	SRAELKDYKA	KNIAGERRLL	LVTSTQGEGE
	70	80	90	100	110	120
	130	140	150	160	170	180
m150.pep	PPKEAVVLHK	LLNGKKAPKL	DKLQFAVLGL	GDSSYPNFCR	QAGKDFDRR	FEELGAKRLLER
a150	PPEEAVVLHK	LLNGKKAPKL	DKLQFAVLGL	GDSSYPNFCR	QAGKDFDRR	FEELGAKRLLER
	130	140	150	160	170	180
	190	200	210	220	230	240
m150.pep	VDADLDFTAS	ANAWTDNIAA	LLKEEAAKNR	ATPAPQTPP	PAGLQTAPD	GRYCKAAPPFAA
a150	VDADLDFAAA	ADGWTDNIAA	LLKEEAAKNR	ATPAPQTPP	PAGLQTAPD	GRYCKADFPFAA
	190	200	210	220	230	240
	250	260	270	280	290	300
m150.pep	LLANQKITAR	QSDKDVRHIE	IDLSGSDLHY	LPGDALGVWF	DNDPALVREI	LDLLGIDPAT
a150	LLANQKITAR	QSDKDVRHIE	IDLSGSDLHY	LPGDALGVWF	DNDPALVREI	LDLLGIDQAT
	250	260	270	280	290	300
	310	320	330	340	350	360
m150.pep	EIQAGGKMMP	VARALSSHFE	LTQNTPAFVK	GYAAFAHYE	ELDKIIADNA	VLQDFVQNTPI
a150	EIQAGGKTLP	VASALLSHFE	LTQNTPAFVK	GYAPFADDDE	LDRIAADNAV	LQGFVQSTPI
	310	320	330	340	350	360
	370	380	390	400	410	420
m150.pep	VDVLHRFPAS	LTAEQFIRLL	RPLAPRLYSI	SSSAQAEVGE	VHLTVGVVRF	EHEGRARTGG
a150	ADVLHRFPAS	LTAEQFIRLL	RPLAPRLYSI	SSSAQAEVGE	VHLTVGVVRF	EHEGRARAGG
	370	380	390	400	410	420
	430	440	450	460	470	480
m150.pep	ASGFLADRLE	EDGTVRVF	VERNDGFRLE	PDSRKPIVMIG	SGTGVPFRAF	VQRAAENAE
a150	ASGFLADRLE	EDGTVRVF	VERNDGFRLE	PDSRKPIVMIG	SGTGVPFRAF	VQRAAENAE
	430	440	450	460	470	480
	490	500	510	520	530	540
m150.pep	GKNWLIFGNP	HFAFDLYQT	EWQFQAKDGF	LHRYPFAWSR	DQEEKIYVQD	KIREQAEGLW
	490	500	510	520	530	540

430

a150	GKNWLF	FGNPHF	ARDFLY	QTEWQQ	FAKDGFL	HRYDFA	WSRDQEE	KIYVQDK	IREQA	EGLW
	490	500	510	520	530	540				
m150.pep	550	560	570	580	590	600				
	QWLQEG	AHIYVC	GDAAKM	AKDVEA	ALLDVII	GAGHLDE	EAGAEY	LDMLRE	EKRYQR	DVYX
a150	QWLQEG	AHIYVC	GDAAKM	AKDVEA	ALLDVII	GAGHLDE	EAGAEY	LDMLRE	EKRYQR	DVYX
	550	560	570	580	590	600				

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 613>:

g151.seq

```

1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCACACGTCG ACCACGGCAA
51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTG
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
251 TGGAGCGCGT TTTGGGGATG GTGGATTGCG TCGTCTTGTT GGTGGACGCA
301 CAGGAAGGTC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAACCGTCCG
401 CCCGTCCGAG CTGGGTATATC GACCAGACTT TCGAGTTGTT CGACAACTTG
451 GGTGCGACCG ACGAGCAGTT GGATTTCCTG ATTGTTTACG CTTCAGGTTT
501 GAGCGGCTTT GCCAAGCTGG AAGAAAcga CGAGAGCAGC GATATGCGCC
551 CGCtggttcgA CACCATCCTA AAATACAcgc ctgCACCGAG CGGCAGCGCG
601 GACGAGCCGC TGCAACTGCA AATTTCCTCA CTCGACTACG ACAACTACAC
651 CGGCCGCCCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
701 AAACCGTTGC CGTGATGAAC CACGAGCAGC AAATCGCCCA AGGCCGCATC
751 AACCAGCTTT TGGGTTTCAA AGGCTTGGA CGCGTGCCGC TTGAAGAAGC
801 CGAAGCCGGC GACATTGTGA TTATTTCCTG TATCGAAGAC ATCGGCATCG
851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCTGCC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTAAACA CCAGCCCGCT
951 CGCAGGTACA GAAGGCAAAT TCGTGACCAG CCGCCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATTGCTGACC AACGTGCCC TGCGCGTGGA AGACACCGCC
1051 GatgCCGACG TGTTCCGCGT ATCcgGGCGC GGCGAACTGC ACCTGACGAT
1101 TTTGCTGGAA AATATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAGC
1151 CGCGCGTCGT GTACCGAGAC ATCGACGGTC AAAAATGCGA ACCTTATGAA
1201 AACCTGACTG TGGACGTACc cgacgacAAC CAAGGCGCGG TAATGGAAGA
1251 ACTCGGCCGC CGCCGTGGCG AACTGACCAA TATGGAAAGC GACGGCAACG
1301 GacgCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
1351 CAAGGCGAAT TCATGACCCT GACGCGCGGC GTCGGGCTGA TGAgccacGT
1401 GTTcgacgac tacgcgcccg tcaAACCCGA TATGCCCGGC CGCCACAACG
1451 GCGTactggt GtcccaAGAG CAGGGCGAGG CGGTTGCTTA CGCCTTGTGG
1501 AATCTTGAAG ACCGCGGCCG TATGTTCTGT TCGCCCAACG ACAAATCTA
1551 CGAAGGTATG ATTATCGGCA TCCACAGCCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTCACCAATA TCCGTGCCAG CGGTACCGAC
1651 GAAGCGGTGC GCCTGACCAC GCCGATCAAA CTGAcgcTGG AAGGCGCGGT
1701 CGAGTTTATC GACGATGACG AGCTGGTGGA AATCACGCCG CAatccatcc
1751 gcctgcgcac gcgttacctG AGCGaattgg aacgccgccg tcaTTTTAAA
1801 AagctgGATT AA

```

This corresponds to the amino acid sequence <SEQ ID 614; ORF 151.ng>:

g151.pep

```

1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51  RGITILAKNT AIDYEGCHIN IVDTPGHADF GGEVERVLGM VDCVLLLVDA
101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESS DMRPLFDITL KYTPAPSGSA
201 DEPLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HEQQIAQGRI
251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DNGRTRLEY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVFD D YAPVKPDMPG RHNGVLVSQE QGEAVAYALW

```

501 NLEDGRGMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRMYL SELERRHFK
 601 KLD*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 615>:

m151.seq
 1 ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
 51 AACACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
 101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
 151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
 201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGGCGAAG
 251 TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
 301 CAGGAAGGCC CGATGCCGCA AACCCTGTTT GTGACCAAAA AAGCCTTGGC
 351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCCG
 401 CTCGTCCGAG CTGGGTATC GACCAAACTT TCGAGCTGTT CGACAACTTG
 451 GCGCGGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTACG CTTCAGGGTT
 501 GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
 551 CGCTGTTCTGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
 601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
 651 CGGCCGCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
 701 AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGCATC
 751 AACGAGCTTT TGGGTTTCAA AGGTTTGGA AGCGTGCCGC TTGAAGAAGC
 801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
 851 GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCCTACC GATGTTGAGC
 901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
 951 GGCGGTACG GAAGGCAAAT TCGTAACCAG CCGCAAATC CGCGACCGCC
 1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
 1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GCGGAGCTGC ACCTGACCAT
 1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
 1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
 1201 AACCTGACCG TGGATGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
 1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
 1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
 1351 CAAGGCGAAT TTATGACCTT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
 1401 GTTCGACGAT TACGCGCCG TCAAACCGA TATGCCCGC CGCCACAACG
 1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTTG
 1501 AATCTGGAAG ACCGCGGCCG TATGTTTCGTA TCGCCCAACG ACAAATCTA
 1551 CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
 1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
 1651 GAAGCCGTTT GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
 1701 TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
 1751 GTCTGCGCAA GCCTTACTTG AGCGAATTGG AACGCCGCCG CCACTTTAAA
 1801 AAGCTGGATT GA

This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:

m151.pep
 1 MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
 51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA
 101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
 151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDLIL KYTPAPSGSA
 201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HDQQIAQGRI
 251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IIGIVTITDK DNPKGLPMLS
 301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
 351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
 401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
 451 QGEFMTLTRG VGLMSHVFDG YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
 501 NLEDGRGMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRKRYL SELERRHFK
 601 KLD*

Computer analysis of this amino acid sequence gave the following results: